Volume 34 OCTOBER, 1939 Number 4

The Journal of Abnormal and Social Psychology

EDITED BY

GORDON W. ALLPORT
HARVARD UNIVERSITY

CONSULTING EDITORS

EDNA HEIDBREDER
Wellesley College

FLOYD H. ALLPORT
Syracuse University

RUTH BENEDICT
Columbia University

HENRY E. GARRETT
Columbia University

JAMES Q. HOLSOPPLE
New Jersey State Hospital

CARNEY LANDIS
New York Psychiatric Institute

HENRY T. MOORE
Skidmore College

DAVID SLIGHT
University of Chicago

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION

EDITORIAL OFFICE: EMERSON HALL, CAMBRIDGE, MASS.

PUBLICATION OFFICES: ALBANY, N. Y., AND
THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO

Entered as second-class matter, June 18, 1923, at the Post Office at Albany, New York

LENGTHS OF CONVERSATIONS: A CONFORMITY SITUATION ANALYZED BY THE TELIC CONTINUUM AND J-CURVE HYPOTHESIS

BY FLOYD H. ALLPORT
Syracuse University
AND
RICHARD S. SOLOMON
The Personnel Institute, Chicago *

The Methods and Assumptions of Conformity Measurement

The subject of conformity, broadly defined, is one which possesses considerable present interest both for social scientists and for psychologists. Organized society would be impossible without great areas of conformity in the behaviors of citizens. The study of this phenomenon is, therefore, one of the basic essentials for our understanding both of societal problems and of the individual human relationships which they involve. The quantitative, scientific description of custom, folkways, traditions, and institutions presents a new field, as yet practically unexplored. And in the making of such studies, the modern method of social psychology may well be brought to bear in rendering more precise and effective some of these earlier sociological conceptions. Such investigations, moreover, may have practical value for the future; for governments are changing in their form, and the increasing regimentation of opinions, attitudes, and behaviors in great conformity patterns seems to be a characteristic of the new order.

However useful they may be for administration and control, the familiar concepts of rule, juristic law, customs, cultural traditions, and mores have proved unsatisfactory for scientific study. They are too implicit, unmeasurable, and vague in their application. In a preceding article, there have been discussed some of the inadequacies of the earlier conformity concepts, and a method has been outlined for more precise and quantitative study.[1] The

* The investigation reported in this article was carried on while Dr. Solomon was associated with the Department of Psychology at Syracuse University.

[1] Allport, Floyd H. Rule and custom as individual variations of behavior distributed upon a continuum of conformity. *Amer. J. Sociol.*, 1939, **44**, 897–921.

present article is an application of this method to the study of a particular conformity situation. As a preparation for what follows, therefore, we shall briefly review the basic features of this method.

The chief defects of the earlier conformity concepts are (1) that they are stated in terms of abstract social norms rather than in terms of actual observable events, (2) that they employ administrative dichotomies rather than continua of measurement, and (3) that they do not afford a method of discovering *conformity* in contradistinction to simple *uniformity*. The method here suggested is an attempt to remedy these defects. As an underlying assumption and working tool, it employs a form of distribution which has been previously found to be a characteristic feature, or indicator, of conformity situations. We refer here to a formulation which is known as the "J-curve hypothesis."[2]

For the understanding of this hypothesis, certain methodological conceptions need to be explained. We must distinguish between several types of continua. There is, first, the *empirical continuum*. This is one of degrees of variables such as are employed in physics—for example, time, space, weight and velocity. This continuum may be applied in societal situations; for instance, in studying the speed with which a motor car passes an intersection, the time of arrival of workers at a factory, or the width of trousers or skirts.

Another form of measurement is that which employs a *personality trait continuum,* a device familiar to psychologists through the use of personality trait tests. It is possible to construe the units of an empirical scale in terms of a personality continuum, as, for example, when the time of arrival of workers at a factory has been proved to have a dependable relationship to *promptness* or *dilatoriness* as traits characteristic of their behavior as individuals; or when the speed of passing an intersection may be taken as an indicator of the characteristic boldness or caution of the individual motorist.

The third type of measurement variable employs what is known as a *telic continuum*. This form of measurement is novel

[2] This hypothesis is explained in detail in the article to which we have referred above. A fuller statement with supporting data is to be found in a still earlier article: see F. H. Allport, The J-curve hypothesis of conforming behavior, *J. soc. Psychol.*, 1934, 5, 141–183. Both these articles may be consulted as background material which will be helpful in connection with the present article. See also M. C. Dickens and R. S. Solomon, The J-curve hypothesis: certain aspects clarified, *Sociometry*, 1938, 1, 277–291.

in psychological work, and had been particularly devised for studies of this character. The question upon which this continuum is based is as follows: "To what degree does the end result of the behavior, or its manner of execution, carry out a known purpose of an individual or individuals who are effective in the situation?" An example would be the degree to which the worker conforms to the purpose of factory managers implicit in the rule prescribing a certain hour for arrival at work. This is a finite continuum with definite ends. The left end is the step of the continuum-scale which represents the *complete fulfillment* of the purpose. In other words, in our illustration, it includes all workers who arrive "on time" (i.e. *on* or *before* the time set). The successive steps on the continuum to the right represent consecutively lesser degrees of fulfillment of the purpose; in other words, in our illustration, successive intervals of "lateness," whose lengths are established by the managers for convenience in administration and discipline. The extreme right end of the continuum is the step which represents the least possible degree of fulfillment of the purpose which is still *some* recognizable fulfillment. In our example it would include the last interval of lateness just before that interval is reached at which the worker may be said to be carrying out *none* of the purpose of a rule for factory opening time.

Another example of a telic continuum is afforded by the behavior of motorists upon approaching a red traffic signal light at an intersection. Those who stop completely would fall on the first step of the continuum, at the left, i.e. the position of complete fulfillment of the purpose of the rule. Those who almost, but not quite, stop would be on the next step to the right. Those who slow up only moderately would be still further to the right on the continuum—and so on, until we come to the category of those who go through the intersection with no change of speed. These latter would lie completely off the continuum to the right.

It will be seen that the empirical, the personality, and the telic continua are related. It is sometimes possible, even valuable, to construe the scale units of one of them in terms of scale units of the others. These continua, however, are distinct in meaning, and must be separately understood and applied in the study of conforming behavior. Where the purpose implicit in the telic

continuum is one of *general acceptance* in a population and is embodied or in some social rule, law, or statement of a "custom," we have what we may call a "societal-telic" continuum. A societal-telic continuum is thus our true *continuum of conformity*.

A further concept necessary for understanding the J-curve hypothesis is that of a "field or situation of conformity." Such a situation is said to exist when there is a generally accepted, though not necessarily explicitly stated, rule relative to a given situation, and when 50 per cent or more of the population fall upon the left-hand step of a telic continuum whose variable is the degree of fulfillment of the purpose implicit in this rule. For example, when motorists approach a traffic signal which shows red, observations have shown that 50 per cent or more of the motorists completely stop. We are not concerned, for the purposes of this definition, with the behavior of those who do not stop. The fact of a generally accepted rule with 50 per cent or more falling on the step of complete fulfillment of the purpose of the rule constitutes a true conformity situation.

We are now ready for the statement of the J-curve hypothesis.

If, in any field of conformity (see definition given above) we apply a scale whose steps are variations of behavior which represent successive recognizable degrees of fulfillment of the "accepted common purpose," ranging from the prescribed or "proper" act, which most completely fulfills the purpose (on the left), to that which gives it the least recognizable amount of fulfillment (upon the right), the following will occur:

a. More instances will fall upon the step at the extreme left than upon any other.

b. The successive steps from left to right will have a respectively diminishing number of instances.

c. The decline in the number of instances will decrease as we proceed by successive steps from left to right.

To state the hypothesis in another form: *Degrees of conformity in a conformity field are distributed upon their appropriate telic (or conformity) continuum in decreasingly diminishing proportions as one proceeds from the fully conforming act or degree of attribute of the act toward the least conforming act or degree of attribute of the act.*

Or, more simply: *In a conformity situation the distribution*

of degrees of conformity upon their appropriate continuum is in the form of a curve of positive acceleration toward the mode.

Concerning the exact rate of acceleration of the slope of the J-curve, in this hypothesis, no general statement is made. This rate will depend upon the situation, the strength of conformity producing influences, and probably other factors.

It will be noted that neither the rule nor the fact of a conformity distribution is posited in advance by this hypothesis. Each is merely hypothesized for experimental purposes. The actual spread of the data gives an adequate and acceptable proof of whether they exist in the situation, and whether the variable of the continuum used correctly states the commonly accepted purpose and rule. Though we adopt as a working tool "a rule" and an experimental, or conformity, continuum, the true fact of a conformity situation is established only by the proportions of the data as they fall upon that continuum. If there is a conformity situation, then the continuum will measure degrees of that conformity; otherwise it will measure degree of fulfillment of a purpose, but not of conformity to a generally accepted rule. What the rule really is and whether there is conformity (as we have defined it) are thus inductively and experimentally determined.

Referring back to our example of the motorists, extended observations have revealed that in this situation the single J form of distribution of our hypothesis actually obtains. That is, the proportions of motorists who carry out the purpose of the rule less and less adequately by lesser and lesser degrees of slowing, decrease in successive diminishing proportions as one goes toward the right extreme of the continuum. The same distribution holds for degree of punctuality in workers arriving at the factory, and for many other fields of behavior which are conformity situations by our definition and by common acceptance.

A second part of the J-curve hypothesis deals with the distributions to be expected from conformity situations when the empirical rather than the telic continuum is used. In such plotting of the data, the resulting distribution usually resembles a double J-curve of asymmetrical character; that is, two unequal J's placed back to back. Since this part of the hypothesis needs further study, we have not made it crucial to our present method,

but have used it only for suggested explanations in terms of certain factors influencing the distribution. In speculating upon the probable influences which have produced this distribution, however, we are led to enumerate four possible factors. These are: (a) conformity-producing agents, who tend to skew the distribution to the left and steepen the mode, (b) biological or physical inertia, which tends to push the distribution toward the right in opposition to the conformity-producing agents, (c) individual differences in ability or personality traits, which produce the tails of the double J-curve, and (d) simple chance factors such as accidents or other unpredictable irregularities.

We now have before us a method through which conformity in a situation can be distinguished from uniformity. The empirical continuum gives us simple uniformity. When the data are translated upon a telic conformity continuum, we can determine whether they yield a true conformity situation as we have defined it. By comparing our distribution with that to be expected by the J-curve hypothesis, we can determine whether the continuum we are using is a correct one for measuring the conformity involved, and, if so, what the variable is upon which the population is conforming. The comparison of empirical, telic, and personality distributions, furthermore, may show us the relation of the conformity pattern to natural individual differences and other aspects of psychological interest.

The specific questions which may be conceived as capable of answer through this method are as follows: (1) Is there uniformity? If so, in what respect? (2) Is there conformity? (3) If so, in what aspect of the behavior does the conformity lie? (4) Are there recognizable degrees of conformity, or only a dichotomy? (5) What proportion of the population fully conform? (6) What proportion conform in a lesser degree? (7) What proportion do not conform at all? (8) What is the relation of individual personality differences to the conformity distribution? (9) What clues are there as to the conformity-producing agents? (10) What degrees of individual variation exist within the range of full conformity? (11) What does the conforming distribution mean in terms of the interests, values, or motives of the population?

For the purpose of testing this method, it was thought desirable to employ, at first, a simple and familiar situation in which

behaviors are easily measured and recorded. The authors therefore chose the behavior situation of conversations between individuals who are working in a library reading room or attending a religious service. The empirical variable chosen was *length* of the conversation. The specific behavior involved, in so far as this was a conformity phenomenon, was the tendency of individuals in these situations to restrict their conversations to a shorter length than in the more informal situations of daily life. In order to secure a contrasting situation for control purposes, the investigation was extended to include the measurement of conversations in the living room of a fraternity house. The library situation was in the main study room of the Syracuse University library. The church situation was religious services at the Temple of Concord at Syracuse, New York (denomination, Jewish Reformed). The living-room situation was in one of the fraternity houses at Syracuse University.

Collection and Presentation of the Data

The method of collecting the data may be best illustrated by describing it for the library situation. The observer sat at one of the long reading tables from which he could gain a fairly good view of the room. When two persons within his view began to converse, he started his stop watch, stopping it as soon as they had finished. The criterion of whether the conversers had finished was their ceasing to talk and starting to do something else. During the timing, the stop watch was held in the observer's pocket so that its click would not be audible at the distance of the conversers. After the watch was stopped, it was unobtrusively taken out of the observer's pocket and the time recorded. The record was made upon a small pad concealed on the table behind a book which the observer pretended to be reading. As soon as one conversation time was recorded, the observer held himself in readiness to time the next one which should come within his field of observation. Only one conversation was observed and recorded at one time; the content and topics of the conversations themselves, as well as all other aspects except that of their duration, were ignored.

On the average the observer was from twenty-five to forty feet away from the pair under observation. He avoided directly facing or looking at them as much as possible; and he kept his

position constant for a considerable period so as not to attract attention. From all outer appearances not more than about five individuals in the library situation seemed to realize that they were being observed or their conversation timed. In the church situation, so far as the observer could judge, none of the individuals showed such awareness. In the fraternity living room about twenty-five subjects betrayed a suspicion that they were being watched, in some instances coming over and asking the observer what he was doing. These latter cases were all omitted from the record.

In the church the observer sat usually toward a corner and recorded conversations which came within his range of observation, for the most part to the side and front. The watch as before was in his pocket and a record pad on his lap. The ending of the conversation was here determined not by the conversers doing something else, but by a readily recognizable change in the position of the head. The conversations heard in the church were uniformly subdued to a whisper, whereas those in the library and clubroom were in a talking tone.

In the fraternity living room the observer sat at the entrance from a small hall from which he could observe people in the middle of the room or seated around the edges upon three of its sides. Most of the observations were made just before or after the dinner hour, a time when there was no urgency of classes and the men were talking about a variety of topics. The subjects were frequently not the only persons in the room, but none of those observed was obscured by the presence or movements of others. The observer held a magazine, pretending to be reading. As before, the stop watch was operated in his pocket, and the record made on a small pad concealed in his lap. A total of two hundred cases was obtained for the church situation, eight hundred for the library and four hundred for the clubroom.

The data for the three situations, in number and percentage of cases in successive ten-second intervals are presented in Tables 1, 2, and 3. The same data have been plotted in percentages in ten-second intervals in Figures 1, 2, and 3, for the three situations respectively.

It will be seen that the continuum employed in these graphs is *empirical*. It represents simply physical units of time, recording a certain aspect of the behavior, not acts of behavior as

TABLE 1

Frequency Distribution in Terms of Ten-Second Intervals of the Duration of Conversation in the Church Situation

N=200

MINUTES	SECONDS	NO. OF CASES	PER CENT
0	0– 9.9	22	11.0
	10–19.9	32	16.0
	20–29.9	55	27.5
	30–39.9	23	11.5
	40–49.9	18	9.0
	50–59.9	12	6.0
1	0– 9.9	9	4.5
	10–19.9	6	3.0
	20–29.9	5	2.5
	30–39.9	4	2.0
	40–49.9	3	1.5
	50–59.9	2	1.0
2	0– 9.9	2	1.0
	10–19.9	1	.5
	20–29.9	1	.5
	30–39.9	0	0.
	40–49.9	1	.5
	50–59.9	1	.5
3	0– 9.9	0	0.
	10–19.9	0	0.
	20–29.9	1	.5
	30–39.9	1	.5
	40–49.9	0	0.
	50–59.9	0	0.
4	0– 9.9	0	0.
	10–19.9	1	.5
	20–29.9	0	0.
	30–39.9	0	0.
	40–49.9	0	0.
	50–59.9	0	0.
		200	100.0

wholes. These time values represent also the talking behavior of two people, rather than the behavior of one person; for they indicate the duration of the entire conversation. Some individuals probably took a relatively shorter part in the conversation

TABLE 2

FREQUENCY DISTRIBUTION OF THE DURATION OF CONVERSATIONS IN TERMS OF TEN-SECOND INTERVALS IN THE LIBRARY SITUATION

N=800

MINUTES	SECONDS	NO. OF CASES	PER CENT
0	0– 9.9	33	4.1
	10–19.9	42	5.1
	20–29.9	26	3.2
	30–39.9	67	8.3
	40–49.9	124	15.5
	50–59.9	93	11.6
1	0– 9.9	72	9.0
	10–19.9	56	7.0
	20–29.9	59	7.4
	30–39.9	55	6.9
	40–49.9	48	5.7
	50–59.9	17	2.6
2	0– 9.9	12	1.5
	10–19.9	14	1.7
	20–29.9	13	1.6
	30–39.9	12	1.5
	40–49.9	7	.87
	50–59.9	6	.75
3	0– 9.9	5	.60
	10–19.9	5	.60
	20–29.9	5	.60
	30–39.9	4	.50
	40–49.9	5	.60
	50–59.9	4	.50
4	0– 9.9	2	.25
	10–19.9	3	.35
	20–29.9	3	.35
	30–39.9	1	.12
	40–49.9	2	.25
	50–59.9	1	.12
5	0– 9.9	0	0.
	10–19.9	0	0.
	20–29.9	2	.25
	30–39.9	0	0.
	40–49.9	1	.12
	50–59.9	0	0.
6	0– 9.9	1	.12
	10–19.9	1	.12
	20–29.9	0	0.
	30–39.9	0	0.
	40–49.9	0	0.
	50–59.9	0	0.
7	0– 9.9	1	.12
		800	99.99

than others. Owing to the large number of cases, however, it is a reasonable assumption that these differences will fairly well cancel and balance one another, so that the conversation times will be on the whole proportional to the individual talking times.

TABLE 3

Frequency Distribution in Terms of Ten-Second Intervals of the Duration of Conversation in the Clubroom Situation

N=400

MINUTES	SECONDS	NO. OF CASES	PER CENT	MINUTES	SECONDS	NO. OF CASES	PER CENT
0	0– 9.9	6	1.50	5	0– 9.9	2	.50
	10–19.9	7	1.75		10–19.9	1	.25
	20–29.9	8	2.00		20–29.9	1	.25
	30–39.9	8	2.00		30–39.9	0	0.
	40–49.9	11	2.75		40–49.9	1	.25
	50–59.9	16	4.00		50–59.9	2	.50
1	0– 9.9	19	4.75	6	0– 9.9	0–	0.
	10–19.9	21	5.25		10–19.9	2	.50
	20–29.9	24	6.00		20–29.9	0	0.
	30–39.9	34	8.50		30–39.9	0	0.
	40–49.9	48	12.00		40–49.9	1	.25
	50–59.9	44	11.00		50–59.9	1	.25
2	0– 9.9	21	5.25	7	0– 9.9	1	.25
	10–19.9	18	4.50		10–19.9	0	0.
	20–29.9	15	3.75		20–29.9	1	.25
	30–39.9	15	3.75		30–39.9	0	0.
	40–49.9	12	3.00		40–49.9	1	.25
	50–59.9	9	2.25		50–59.9	0	0.
3	0– 9.9	6	1.50	8	0– 9.9	1	.25
	10–19.9	4	1.00		10–19.9	0	0.
	20–29.9	4	1.00		20–29.9	0	0.
	30–39.9	4	1.00		30–39.9	0	0.
	40–49.9	3	.75		40–49.9	1	.25
	50–59.9	4	1.00		50–59.9	0	0.
4	0– 9.9	3	.75			400	100.0
	10–19.9	5	1.25				
	20–29.9	5	1.25				
	30–39.9	4	1.00				
	40–49.9	3	.75				
	50–59.9	3	.75				

It should also be remembered that since it takes two to make a conversation, each individual is responsible for carrying it on. For our purpose, it is as useful to consider the behavior to be concerned with the length of time of "carrying on a conversation" as to consider it as the "length of time of talking."

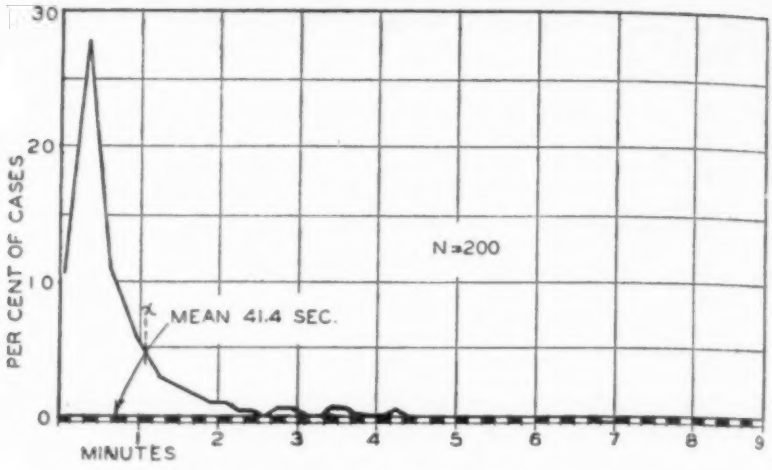

FIG. 1. DISTRIBUTION OF DURATION OF CONVERSATIONS
Church situation. Empirical continuum (time).

When we examine the data as presented in the tables and graphs, we see unquestionable evidences of uniformity. The greatest degree of uniformity seems to be in the church situation, the next largest is in the library, and the smallest in the clubroom. The church conversations range in duration from 10 seconds (or under) to 4 minutes and 10 to 19 seconds, with a sharp mode of 27.5 per cent in the interval 20 to 29 seconds.

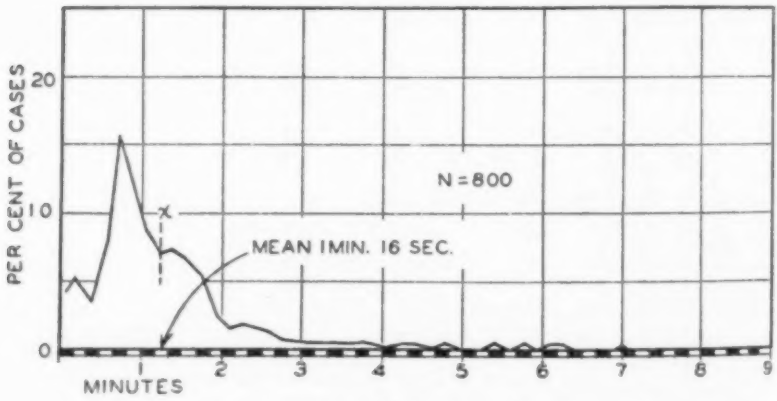

FIG. 2. DISTRIBUTION OF DURATION OF CONVERSATIONS
Library situation. Empirical continuum (time).

From this point it descends steeply to right and left, with slight irregularities on the right. In the library, conversations range from 10 seconds (or under) to 7 minutes and 0 to 9 seconds, with a mode, not quite so sharp as that of the church (15.5 per cent), at 40 to 49 seconds. From here the curve falls steeply to right and left, although not so steeply as in the church and with slightly larger irregularities. In the club the conversations range from 10 seconds (or under) to 8 minutes and 40 to 49 seconds. The mode here is still lower (12 per cent) and falls at 1 minute and 40 to 49 seconds. This means that the modal conversation of the clubroom was about a minute longer than

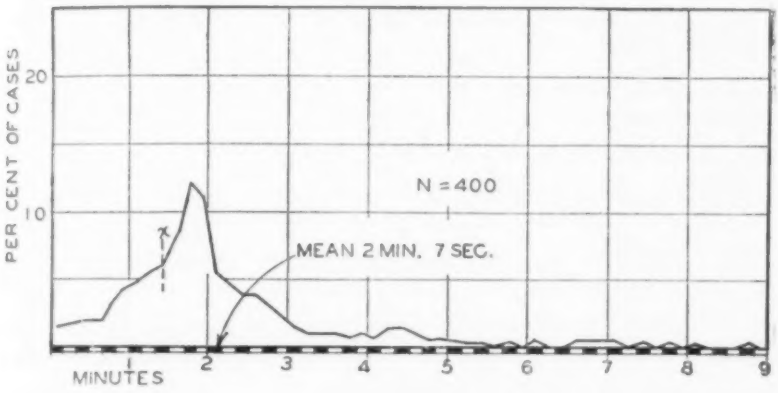

Fig. 3. Distribution of Duration of Conversations
Clubroom situation. Empirical continuum (time).

that of the library and about one and one-third minutes longer than that of the church. The slopes fell away from the clubroom mode with considerable irregularity toward the extremes; and unlike the other two distributions, positive acceleration is lost on approaching the mode from the right. Uniformity is greater in church and library than in the clubroom. The mean for the church distribution is 41.4 seconds, for the library 1 minute 16 seconds, and for the clubroom 2 minutes 7 seconds. On account of the sharpness of the mode, none of these distributions resembles the normal frequency curve. In two of them—namely, those for the church and for the library—the curves with some exceptions are in general of the double J-curve type, with positive acceleration on both sides toward the mode.

(The vertical dotted lines, labeled x in these graphs, may be ignored for the present.)

What do these raw data in themselves, when plotted upon an empirical continuum, tell us about the conformity situation? In the first place they are all important, in the sense that without them no other exact knowledge of the situation in respect to this variable could be gained. They are far more reliable than the impressionistic categories of "rule" or "custom." We can also infer that there must be certain psychological factors at work which differ in the three situations employed. Beyond these facts and inferences, however, the knowledge given by our data does not extend. We know what the modal interval of length of talking is, and the range, but we have no proof beyond mere assumption as to what the norm or "accepted" range of conversation length may be. We do not even know that there *is,* in these situations, an accepted norm or rule. Nor do we know definitely the purpose or meaning of the behavior underlying whatever conformity might be found to exist. All we have is the lengths of time which the individuals talked.

Suppose we should state that the "custom" with respect to talking in these three situations is within the range of about four minutes for the church, seven minutes for the library, and nine minutes for the clubroom. Are these points to be taken as true definitions of the custom range, or only negative statements that such a range cannot fall outside these figures? Are all the individuals who converse within the four, seven, and nine-minute limits, respectively, conforming to the custom? There is no way of answering this question. We might suspect that a custom against long conversations does exist in the church and library, but *not* in the clubroom; and that the actual data would demonstrate this fact. Here now are the actual data—but can we decide? What criterion can we apply to our data, to distinguish conformity from non-conformity situations? We cannot fall back on the mode as the true index of our "custom," for, although the modes are sharp, in no case do they include over 28 per cent of the population—too small a proportion, clearly, to signify conformity. Or suppose that we say the "custom" embraces the mode plus all the individuals to its left as true conformers. Though this hypothesis seems more plausible, it is

still an arbitrary procedure, for it derives the "custom" from the lay of the statistical majority without ascertaining whether that majority is conforming to a norm or merely exhibiting natural human similarities of behavior. Essential as the data of observation are, in themselves they are not enough to establish the fact of a conformity situation. They give us a picture of uniformity, but not of conformity.

The Need for a Hypothetical Conformity Continuum and Its Relation to J-Curve Logic

The problem of whether a conformity situation exists in our data, and, if so, in what aspect the conformity lies, can be solved by a definite procedure. We can construct a hypothetical continuum of conformity, and apply it to our data. Any one of a number of variables could be chosen for this experiment. We could, for example, use a continuum of propriety or decorum to be observed in public places, or a continuum of conduciveness of the behavior to studiousness or to religious worship Suppose then, we shall have constructed such a continuum upon which our raw time data can be replotted. How shall we know that it is the right continuum, that is, that we shall have discovered where the true conformity lies.

The answer to this question is in part self-evident. For we can see whether a conformity field results—whether 50 per cent or more of the cases fall upon the first step (complete fulfillment) of this continuum. If we find that such a phenomenon occurs, we can infer that at least half the population is fully conforming to a rule which expresses a certain known purpose. But we can go further than this. The J-curve hypothesis gives us a pattern of distribution apparently characteristic of the known situations so far studied which are conformity situations by common acceptance and understanding. Following this hypothesis we took a rule and purpose which were known to us as a matter of common knowledge (in law, economic regulations, etc.). We then constructed a telic conformity continuum based upon the degree of fulfillment of this same purpose as its variable. The experimental finding, when the data were replotted on this continuum showed that, in some cases (most of those studied), a conformity situation (according to our definition) existed—that is, 50 per cent of the population fell on the first step of the telic

continuum. Whenever this occurred, the single J form of the *entire* distribution resulted.

We can now reverse the logic of this procedure. Without *knowing* definitely the rule or its purpose, we can *assume* one for experimental purposes, and construct upon this purpose a trial conformity continuum. If upon replotting the data upon such a trial continuum, we find that 50 per cent or more fall upon its first step and the entire distribution takes the J-form, we can then conclude (1) that a conformity situation exists, and (2) that our trial continuum is the correct one—in other words, that its variable is the one on which the conformity of the situation is based. It shows, in short, where that conformity lies. There might, of course, be other, yet unknown, continua which would also reflect some conformity in the situation; but we *know* at least that this one does.

THE CONSTRUCTION OF A HYPOTHETICAL CONTINUUM OF CONFORMITY

How then shall we construct a trial, or hypothetical, conformity continuum? The first step in building any continuum is the selection of a variable. This variable cannot be pure conformity as such, for that does not have any meaning. To have conformity there must be something which people are conforming in or about. Placing ourselves in the position of students in the reading room of a library, the writers assume that the most probable explanation of the shortening of conversations here as compared, for example, with the clubroom situation, is the avoidance of annoyance to others who wish to read or study. The very students who converse are also those who, on other occasions, wish to study and therefore to have *other* students shorten their conversations. So they become in their turn conformity-producing agents (through expressions of disapproval) as well as participants. The psychological considerations underlying a conformity continuum upon this variable would, therefore, have the advantage of being meaningful both to the participants in the practice of shortening conversations and to the conformity-producing agents. The considerations here advanced seem to us to apply equally well to the church and to all similar situations. Hence it seemed advisable to construct our hypo-

thetical conformity continuum upon the basis of avoidance of annoyance.[3]

How, then, can we proceed to develop a scale upon which can be measured the degree of avoidance of annoyance due to length of conversation? We might ask individuals directly how much annoyance they thought would be occasioned by conversations of different lengths. This method, however, would presuppose a degree of experience in judging which most persons would probably not have acquired. If carried out under exact laboratory conditions, moreover, it might be inapplicable to the present practical setting. A more reliable way seemed to be to judge the degree of annoyance in terms of definite acts which individuals perform when annoyed. We can then try to determine the time, in the duration of the conversation, at which these same acts would be likely to be performed; and finally, through the intermediation of the acts, derive a scale of equal annoyance units which can be stated in terms of their equivalent time values. This method has an additional value in that the acts which people commit when annoyed by conversation are often the direct agencies for influencing the conversers to limit their talking.

Our first step, therefore, was to select certain acts or readily described introspective states through which annoyance seems to be expressed. According to accepted principles of scale construction these "landmarks" should be definite, clearly distinguishable categories, characteristic in the situation and readily applicable in practice. They should also cover fairly well the full gamut of degrees of the variable (annoyance) from the least to the greatest. The following five general categories of act or state were finally selected: (a) feeling slightly distracted from what one is doing, (b) beginning to notice the conversation, (c) experiencing a tendency to stare at the individuals, (d) a tendency to ask them to stop talking, and (e) a tendency to have someone in authority make them stop talking.

For the scaling of these items a modification of the Thurstone

[3] Annoyance could, of course, be occasioned by loudness of conversation as well as by length. A hypothetical continuum of conformity could as logically be constructed upon this variable. This fact, however, does not hamper our present investigation, since the raters in making the annoyance scales were instructed to imagine conversation of an intensity of sound constant within each situation. It is probable also that in the conversations observed the tone was fairly uniform because of the restrictions which the situation imposed.

technique was employed.[4] Since so few items or statements were to be scaled, Thurstone's method of throwing them into eleven equal piles was replaced by the method of graphic placement of the statements along a line according to their estimated position between the two extremes of the variable. Scale value positions were then determined in the same way as in Thurstone's psychophysical procedure.

Fifty-six individuals, all of them students in an introductory class in psychology in Syracuse University, were employed as raters for making the scale. The statements of the acts or states were given, in mimeographed form, in single lines which were to be clipped out, so that the strips could be pasted vertically along the edge of a standard sheet of paper according to the position in which the rater judged they should fall. Both oral and written directions were given for this work. Three sets of the statements were given to each rater. With each list of statements and its sheet there was stated the situation to which the rater should consider his judgments as applying. The first was that of two people conversing in church, synagogue, or chapel during services. The second was two students conversing in the main study room of the University library. The third was two people conversing in a clubroom, fraternity, or sorority living room. The list of statements as given to the raters for the library situation is reproduced below.

LIBRARY

I feel like stopping my work and staring at the individuals or actually do so.
I begin to notice the conversation.
I feel like getting up and asking them to stop or actually do so.
I am slightly distracted from what I am doing.
I feel like asking the librarian to make them stop or actually do so.[5]

For the other situations there were slight changes of wording in order to make the statements appropriate. For the word "librarian" there was substituted, for example, "minister" in the church situation and "president" for the club. "Distracted

[4] Thurstone, L. L., and Chave, E. J. *The measurement of attitude*. Chicago: University of Chicago Press, 1929.

[5] It will be noted that three of the statements give the alternatives of either "feeling like" doing an act or actually doing it. While it was recognized that this form might increase ambiguity in rating, it was felt that a greater error, due to individual differences in personality traits among the raters—such, for example, as ascendance—would be avoided. The emphasis is thus thrown upon a degree of annoyance which they feel merits the inhibitory act stated, rather than the requirement as to who is to do the act.

from the activities of worship" was substituted in the church list for the phrase "from what I am doing" as used in the library situation. The statements were presented in random order, and in a different order for each of the situations. The distance from the left-hand extreme, where the statement representing the least annoyance was to be pasted, to the right-hand edge of the sheet was 9.5 inches. The ratings were not performed in the classroom, but by the students working privately, so as to avoid any effects of social facilitation or conformity attitudes in the rating.

When the ratings had been completed, the sheets were assembled and subjected to the following treatment. The point where the line of type describing the statement fell along the bottom edge of the sheet was measured in half-inch units from left to right; and for each statement in each situation the unit where each rater placed that statement was noted on a tabulation sheet. The frequencies of placement of the statement in each of these nineteen half-inch intervals were then totaled for each of the five statements in each of the situations, and the percentages calculated for each. For each statement in each situation separately, a graph of cumulative proportions was made. A vertical line was then dropped from the fiftieth percentile (median) value to the base line. Verticals were similarly dropped from the twenty-fifth and seventy-fifth percentile. The intersection of the median with the base line was taken as the scale value of the statement. The interquartile ranges (Q-values) delimited by the twenty-fifth and seventy-fifth percentile were recorded as the index of ambiguity for that statement. The median and Q-values, however, were recorded not in terms of the original nineteen half-inch units, but in inches and decimals of inches along the base line from the left-hand extreme.

Table 4 presents these scale values and Q-values. We now have our acts or states expressing annoyance scaled on a continuum of the degree of annoyance which they express. Although our "landmark" acts are few and not equally placed, they are probably adequate for our purpose, since we are looking rather for relative proportions in the general shape of the final curve than for distributions in terms of small degrees. *Three* being the approximate Q-value which, on a scale of eleven steps, is usually taken as the maximum index of ambiguity permitted for a scaled statement, this value should be apportioned to fit

our scale (which is 9.5 inches in length). Three-elevenths of 9.5 inches is 2.6 inches; hence this latter figure is taken as our index of ambiguity. All but two of the fifteen Q-values shown in the table are below 2.6. The exceptions are only slightly above this figure: 3.1 (for statement C in the clubroom) and 2.7 (for statement C in the church). All the others are 2.0 or below. In view of these relatively low Q-values, it may be considered that the scale values now given to our statements of acts are sufficiently reliable for our purpose.

TABLE 4

SCALE VALUE AND INTERQUARTILE RANGES (Q) OF THE STATEMENTS ON A DEGREE-OF-ANNOYANCE CONTINUUM FOR THE THREE SITUATIONS, IN TERMS OF INCHES

(Range of scale is from 0 to 9.5 inches)

STATEMENTS		SITUATION		
		CHURCH	LIBRARY	CLUBROOM
A. Slightly distracted from what I am doing.	S	.70	.175	.175
	Q	1.80	.800	1.400
B. I begin to notice the conversation.	S	1.45	1.25	1.25
	Q	2.00	1.60	1.70
C. Feel like stopping my activities and staring, or actually do so.	S	3.60	3.65	3.90
	Q	2.70	2.00	3.10
D. Feel like getting up and asking them to stop, or actually do so.	S	7.0	7.25	7.45
	Q	2.0	1.40	1.50
E. Feel like asking an authority to make them stop, or actually do so.	S	8.80	8.85	9.05
	Q	1.20	1.10	.90

It is to be noted that statement B ("I begin to notice the conversation") was evidently understood to mean "I begin to notice what is said." This may explain why it is placed second instead of first. Statement A ("I am slightly distracted from what I am doing") has the lowest scale value in all situations, while E ("I feel like asking an authority to make them stop, or do so") is located at the upper extreme of the scale. This latter statement has a scale value of 8.80 for the church, 8.85 for the library, and 9.05 for the clubroom. The scale values of the statements are in the same order in each situation. The three scales for annoyance

value of the acts or states are shown in Figure 4. It is the upper of the two scales presented for each situation.

The degree of annoyance having been determined for each of the expressive acts or states, our next task is to devise a method of getting these degrees of annoyance translated into the units of time at which they would occur in the actual situation as the conversation proceeds. This requirement may be fulfilled through the intermediation of the same acts by means of which the degrees of annoyance were obtained. We need to have our raters employ a second continuum, which will be one of time. They are now to place the acts or states along that continuum at the point of time when they judge they would be likely to do or feel them as the conversation proceeds. Each individual makes the judgment wholly for himself.

To increase the accuracy of this procedure the rater was presented with a chart bearing five parallel, horizontal lines, all slightly over eleven inches in length. Just above each line there was stated one of the five acts or states. Each line was divided into uniform units by vertical divisions showing fifteen-second intervals, with longer division indicating successive minutes up to eight, the number of the minutes being marked below. The rater's task was to check on each line his estimate of the time which the conversation would have to endure before the point would be reached at which he would do the act or experience the state in question. Should the time estimated run over eight minutes, he was asked to write it on the space provided over the right extreme of the line.

Three such charts were given the raters, one for each of the situations. The rater was asked to imagine a degree of loudness of the conversation characteristic of the particular situation. Specifically, in the church situation he was asked to imagine the conversations conducted in a whisper, in the library situation in a tone a little above a whisper, and in the clubroom situation in a low conversational tone. These directions were presented both orally and in mimeographed form as before. Fifty-six raters were used in this experiment, the same, with few exceptions, as those who had been employed for the previous ratings. Over a month, however, had elapsed since they had given the earlier ratings; and the statements were presented in a random order different from that of their previous presentation.

When these ratings were collected, the frequencies of the time-checking on each statement for each situation were tabulated in terms of .5-minute intervals. The median of the distribution of these time ratings for each statement (in terms of minutes and decimals of minutes) was recorded directly as the scale value of that statement. The medians (scale values) and Q-values (interquartile ranges) of the statements are shown in Table 5. Taking

TABLE 5

SCALE VALUES (MEDIANS) AND INTERQUARTILE RANGES (Q) OF THE STATEMENTS ON A "TIME-OF-PERFORMANCE" CONTINUUM FOR THE THREE SITUATIONS, IN TERMS OF MINUTES

STATEMENTS		SITUATION		
		CHURCH	LIBRARY	CLUBROOM
A. Slightly distracted from what I am doing.	Median	1.12	1.25	1.40
	Q	1.70	1.90	1.94
B. I begin to notice the conversation.	Median	1.33	1.58	2.10
	Q	1.7	2.30	2.80
C. Feel like stopping my activities and staring, or actually do so.	Median	2.41	3.15	3.16
	Q	2.30	3.00	2.60
D. Feel like getting up and asking them to stop, or actually do so.	Median	5.40	5.75	5.75
	Q	4.70	3.50	3.70
E. Feel like asking an authority to make them stop, or actually do so.	Median	7.75	7.90	8.10
	Q	3.90	3.20	4.10

as our index of ambiguity three-elevenths of the range from zero time to the longest time checked for any of the statements in a given situation, we find an index of 2.9 minutes for church and library, and 3 minutes for the clubroom. Seven of the 15 Q values, mainly those for the statements with higher scale values, are above their proper ambiguity indices. Three of these seven, those for the library, are not more than .6 minute above the ambiguity index. The others are somewhat higher, two being between 3.5 and 4.0 (3.7 and 3.9), and two being above 4. For the church situation, the Q value of statement D (getting up and asking them to stop or feeling like it) is 4.7.

These upper scale values, therefore, must be treated as less reliable than the others, and as less reliable than the values for

degree of annoyance expressed, found in our previous experiment. This fact, however, does not invalidate the scale for our purpose, since the only acts seriously affected are D and E, which come so far out toward the end of the time range that few observed cases fall upon them. The scale values of these statements could be shifted considerably in either direction without changing our time values for our final annoyance continuum to a degree which

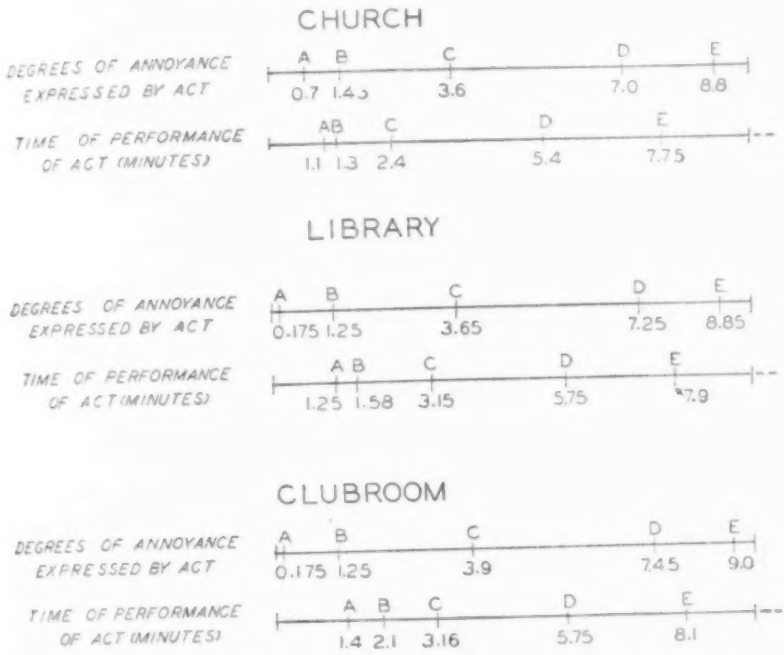
CHURCH
DEGREES OF ANNOYANCE EXPRESSED BY ACT
A B C D E
0.7 1.45 3.6 7.0 8.8
TIME OF PERFORMANCE OF ACT (MINUTES)
AB C D E
1.1 1.3 2.4 5.4 7.75
LIBRARY
DEGREES OF ANNOYANCE EXPRESSED BY ACT
A B C D E
0.175 1.25 3.65 7.25 8.85
TIME OF PERFORMANCE OF ACT (MINUTES)
A B C D E
1.25 1.58 3.15 5.75 7.9
CLUBROOM
DEGREES OF ANNOYANCE EXPRESSED BY ACT
A B C D E
0.175 1.25 3.9 7.45 9.0
TIME OF PERFORMANCE OF ACT (MINUTES)
A B C D E
1.4 2.1 3.16 5.75 8.1

Fig. 4. Graphic Scales for Degrees of Annoyance and Time of Performance of the Acts, for the Three Situations

A=Slightly distracted from what I am doing
B=I begin to notice the conversation
C=Feel like stopping my activities and staring, or actually do so
D=Feel like getting up and asking them to stop, or actually do so
E=Feel like asking an authority to make them stop, or actually do so

would make a difference of more than a very few cases in the plotting upon that continuum. The proportions here are so low that changes in them do not materially alter the shape of the resulting curve. In the church situation it will later be seen that *no cases* fall in the intervals of annoyance in which acts D and E occur.

(See Table 6.) Our "time of performance" scale may, therefore, be considered as sufficiently reliable for our present use.

These scales showing the time at which the act would be performed or the states experienced are presented graphically in Figure 4, as the lower scale for each of the three situations. The statements of the second continuum follow each other, as logic would demand, in the same order as they do in the first. There is also a roughly comparable relative distance between them in

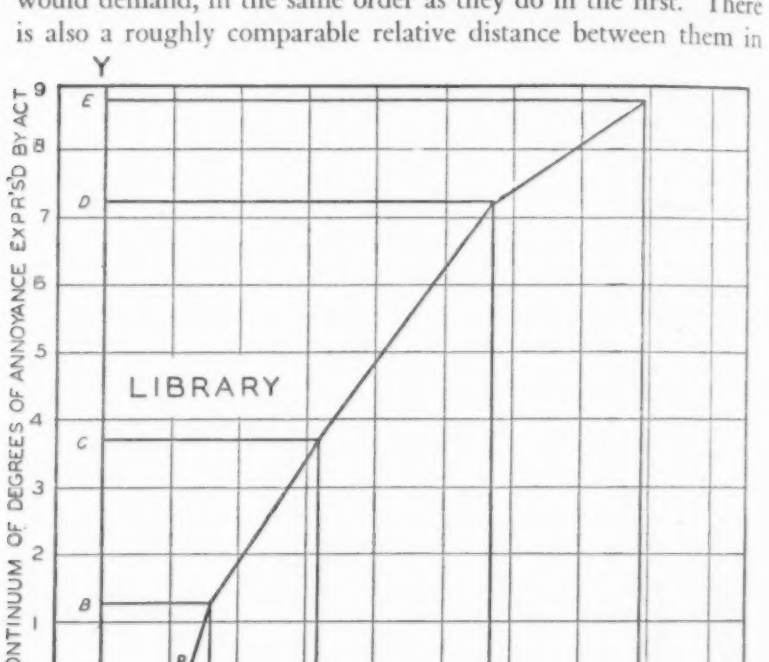

FIG. 5A. CURVE DESCRIBING RELATION BETWEEN DEGREE OF ANNOYANCE AND TIME OF PERFORMANCE OF THE FIVE ACTS (LIBRARY SITUATION)

the two continua for the various situations. The continuum for time of performance, however, shows a decided shortening at both ends.

We now have before us two scales, each giving the scale values of the same set of acts upon two different variables. The first scale variable measures the degree of annoyance which the acts express, the second the length of time of conversation which

would elapse before an act would be committed. We now wish to construct a scale which will show not merely the annoyance value and time value for the particular acts used, but the time value of a whole series of equal annoyance units. We need, in other words, a true scale of annoyance units whose time values are known.

This objective may be reached by finding the curve expressing

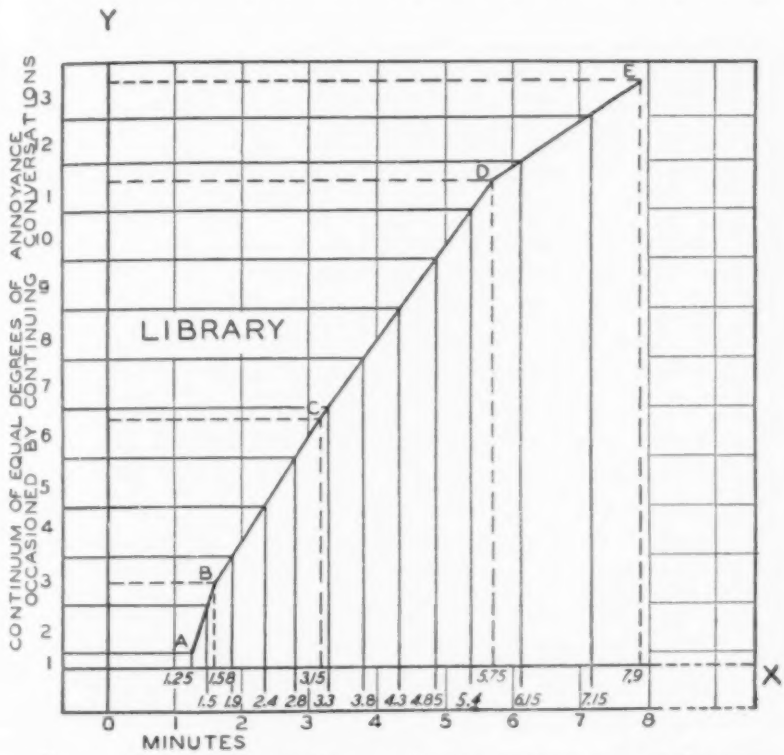

Fig. 5B. Values of Equal Units of Annoyance Occasioned by Conversation (Y-axis) Expressed in Terms of Minutes during Which the Conversation Lasts (X-axis) (Library Situation)

the relationship between the degrees of annoyance expressed by the acts on our annoyance scale and the time at which these acts would be performed. Figures 5A and 5B show how this task is accomplished. If we let the Y-axis (Figure 5A) represent the degree of annoyance and the X-axis the time of performance, we can locate the values for each of the five acts at its proper

point on the respective axes. Thus, in the *library situation,* the value for act A (see Figure 4) is .175 on the continuum for degree of annoyance and 1.25 on the continuum for time of performance. A, therefore, has a Y-value of .175 and an X-value of 1.25. Drawing, in Figure 5A, the horizontal from the proper point on the Y-axis and the vertical from the proper point on the X-axis, we can locate the point of their intersection, p. The values of acts B, C, D, and E are located on ordinate and abscissa in like manner and their points of intersection found. The joining of these points of intersection gives us the "curve" of the relation between our X and Y values, that is, between degree of annoyance and time of performance. This curve is, of course, approximate since the acts on the basis of which it is plotted are only five in number. It is reasonable to assume, however, that if intermediate acts expressive of annoyance could be recognized, measured, and plotted, they would transect ordinate values in such a way that a true curve—rather than a succession of straight lines (as in Figure 5A)—would result, and that this curve would give abscissa values not markedly different from values to be projected from the simplified curve in Figure 5A. The lengths of the interpolated values to be found on the base line would, therefore, bear a ratio to one another similar to that resulting from our present equation line; and the general shape of the curve resulting from a regrouping of the raw data for plotting on the new continuum would not be altered. Figures 5A and 5B have been presented to illustrate the method from the scale values in the library situation. Similar graphs were drawn to determine the line of relationship between degree of annoyance and time of act-performance for the other two situations.

The significance of this procedure should now be clear. We wish to construct a continuum of degrees of annoyance upon which our raw time-data can be regrouped and plotted—a continuum, that is, which will show the frequencies of the population with reference to their avoidance of giving annoyance, in different degrees, through length of conversation. To do so with any accuracy we need an annoyance continuum having a considerable number of step-intervals which will be *equal.* Up to this point it has not been possible for us to define and delimit

a sufficient number of equal units of annoyance on our continuum, since the landmarks, A, B, C, D, and E, are both too few and spaced in too irregular a manner in their annoyance and performance-time values.

By the aid of Figure 5A, however, equal intervals of annoyance can be laid off at will along the Y-axis. It is not necessary for us to be able to recognize or name acts which would express these equal annoyance units. Having laid off these intervals, we project them to the curve describing the relationship between the X- and Y-axes, and drop, from their intersections with that curve, verticals to the X-axis. The point of intersection of these verticals with the X-axis will give us time values corresponding to the boundaries of our successive equal intervals of annoyance. We are thus able to discover the time span of each of the successive equal annoyance intervals. We have, in other words, evaluated a complete scale of annoyance units in terms of the time intervals to which they correspond.[6]

This procedure is shown graphically in Figure 5B, again for the *library* situation. For convenience we have laid off eleven equal intervals (labelled 2 to 12 inclusive) along our continuum of degrees of annoyance (Y-axis), and have determined their intersection, through the curve of relationship of X and Y, by dropping verticals upon the X-axis. These points of intersection upon X give us the time-value limits of the successive equal annoyance intervals laid off on Y. (The time position of the five original "landmarks" is indicated by the intersections of the dotted lines upon X.) If we look at the successive time intervals for the annoyance units in our figure, we shall see that they increase from left to right. This fact suggests the possible operation of the Weber-Fechner law. The time values for the equal annoyance units of the church and clubroom situations were found by a similar graphic procedure. It was necessary to extrapolate one additional annoyance unit for the clubroom situation, the time range of observed conversations being greater than for the church or library. The units of annoyance with their respective time interval values are shown in the first two columns under

[6] This method should be found useful for deriving true telic scales of conformity, with empirical equivalents, in all situations where the observable behaviors can be found which mediate between empirical values and telic values; that is, where acts can be described which can be measured upon both empirical and telic continua.

each situation in Table 6. These time values, for convenience in re-grouping the raw data, have been transtated from minutes (and tenths) to minutes and seconds. We now have a scale of equal units of annoyance experienced through conversations as they continue in time, expressed in terms of the range of time length for each unit.

A further point about Figure 5B, however, requires explanation —the meaning of the short interval at the bottom of the Y-axis (labelled 1 in the figure). It will be noted that, in laying off our equal annoyance intervals, we begin with the Y-value of act A. The reason for this procedure is a logical one. The description of act A reads, "I am slightly distracted from what I am doing." Between the degree of annoyance expressed by this act (or state) and no felt annoyance at all, we have no recognizable landmark. The curve of the relationship between X and Y might be carried out to zero; but its exact shape in this region is not known. Since it seems to become almost horizontal at this point, any increase in time value from zero up to 1.25 minutes (the time value of A) would be accompanied by an increase of annoyance which is so small that it would probably be imperceptible. Annoyance values probably mean very little in the region of the Y-axis below A. We might, in fact, disregard them and say that no annoyance of any consequence is felt unless a conversation lasts at least as long as the time value of act A (1.25 minutes). Qualitative considerations (the meaning of language) also support this view. Annoyance which would be felt in less degree than that indicated by A ("I am slightly distracted from what I am doing") would indeed be very slight. Probably distraction does not fall much below the level described by the word "slightly" until it passes below the threshold and becomes "none," or zero distraction. "*Very* slightly distracted," for example, would be a degree which we probably would not discriminate in practice from "slightly distracted." Now if the distraction is the reason for the annoyance, as seems probable, we may conclude that where distraction becomes unnoticeable, annoyance too would have a zero value. For these reasons it seems logical to exclude the region of Y below A from the annoyance continuum, and begin laying off our units of annoyance where annoyance begins unmistakably to be felt. Conversations of a shorter duration than 1.25 minutes (in the library situation)

may, therefore, be regarded as productive of no annoyance. This fact will have a further significance in our later discussion.

Let us now pause and look at our new annoyance scale from the standpoint of its operation. Disregarding for the moment the interval marked 1 on the Y-axis in Figure 5B and its corresponding time values in the library situation in Table 6, we have on the Y-axis of the figure a series of equal units of felt annoyance, and in Table 6 (under "library") the corresponding, but varying, time values of these annoyance units. They have been derived, it will be remembered, by raters who took the position, not of the conversers, but of others in the vicinity of the conversers who might be disturbed by the conversations. Such persons are the actual or potential conformity-producing agents; for it is out of consideration for their equanimity, or to avoid the disapproval which long conversations may elicit from them, that the conversers limit their time of talking. Let us look at unit 2 of annoyance (Y-axis of Figure 5B) and then at its time interval as shown in the library section of Table 6. These figures mean that when the conversers have talked from 0 time to 1 minute and 16 seconds (where A occurs on the time of performance scale), our unit of annoyance marked 2 will begin to be felt by the one who is reading or studying in the vicinity. (Though this unit is marked 2, it is really the *first* unit of annoyance felt from the situation, as was previously explained. The reason for labelling it 2 will be presently explained.) This degree or unit of annoyance will continue to be felt until the conversers have talked for 1 minute and 30 seconds. While they are talking from 1 minute 31 seconds to 1 minute 54 seconds, should they continue that long, the one being disturbed, having already experienced the first unit of annoyance, will now suffer the second unit. And so on, according to the values in the table for as long as the conversation continues (up to the limit of our ascertained scale values). We may understand in like fashion the meaning of the annoyance units and their time ranges for the other situations.

Our new annoyance scale thus far has been considered as expressing a psychological condition solely in the person *not* conversing—in short, in the conformity-producing agents. Our scale, strictly speaking, is upon neither a telic nor a personality continuum, as these terms are defined earlier in this article, but

is constructed upon a continuum of intensities of the feeling of distraction or thwarting caused by the continued duration of a thwarting stimulus. It is perhaps a continuum of affect, or of the sensory qualities which make up the experience of emotion. The time values merely serve as practical indices, so that we may know when different intensities of the affect-emotion experience are being aroused.

But now let us look at the annoyance continuum from a different standpoint, namely, that of the *talkers*. If our first unit of annoyance, as the affect or emotion of one disturbed, is that which is caused by the talkers speaking for 1 minute and 16 seconds or longer (library situation, unit 2), then, by cutting short their conversation before this time has elapsed, the talkers can avoid giving that degree of annoyance to the other. Similarly by avoiding conversations longer than 1 minute and 30 seconds, they can avoid causing the next unit of annoyance, and so on. Conversely, to talk for longer than 1 minute and 30 seconds is to show that one is not averse to giving annoyance to others to a degree indicated by unit 3 of the annoyance continuum. The talkers thus consider the interests of the non-talkers in these degrees. The degree to which they limit their conversation measures directly the next unit of annoyance which they avoid giving by not continuing into the next time interval.

Clearly, we are dealing now with purpose. A scale which was merely a continuum of psychological states of affect-emotion in those not conversing, becomes for those who *are* conversing a true *telic continuum*. Its variable is the degree to which they fulfill the purpose of not annoying others through the length of their conversation. We do not, of course, imply that the conversers are conscious of these divisions of units of annoyance given to others as they talk. What they probably often are conscious of is a steadily increasing sense that the annoyance they may be causing others is increasing. Our units are simply a refinement of this experience for relating it to overt behavior. Consciousness of the units themselves is necessary neither for the purpose of compliance with the non-talkers' interest, nor for the measurement of the degree of fulfillment of that purpose.

It is clear, then, that our annoyance continuum is really a *telic* continuum of annoyance-giving avoidance so far as the talkers

are concerned; and it is the talkers whose behavior is our primary consideration. But let us go a little further. The experiences of the talkers we have been discussing, and the relevance of their behavior to this continuum, are not solitary events. The facts are highly similar, if not identical, for all talkers in the situation. They are probably all aware, moreover, that the conditions are uniform and that a similar expectancy of avoiding annoyance exists with reference to all conversers. Our telic continuum, thereupon, meets all *a priori* requirements and may be regarded as a true *continuum of conformity* upon the variable selected. It will be recalled from our earlier discussion of telic continua that the first step of such a continuum represents not some lesser degree of fulfillment of the purpose variable, but complete fulfillment. It would not be correct, therefore, for our telic avoidance-of-annoyance-giving scale to begin with the first unit of annoyance given. It should begin with complete fulfillment of the purpose of avoiding the giving of annoyance, that is, with a unit representing the time range of conversing during which *no* annoyance would be felt. The meaning of the small interval of the Y-axis (Figure 5B) below act A is now clear. This, as was explained above, is the region of non-annoyance. Hence it becomes the first step, the step of complete fulfillment of the purpose, on our telic continuum. Regarding this continuum as a continuum of conformity, we see that it becomes the step of complete conformity with the accepted purpose of limiting conversations to a length which would not give annoyance to others.[7] It is therefore entered, with its proper time-interval values, as the first step in the conformity continua of annoyance-avoidance, for the three situations, in Table 6. Since, however, this step represents not a unit of annoyance but absence of annoyance, its number, 1, is enclosed in parentheses.

We shall now think of our annoyance continuum, therefore, as a continuum of conformity. The step intervals are in terms of units of annoyance to others to be avoided through shortening the conversation. Its units are units of annoyance. *It is not a*

[7] The fact that this step is not equal on the graph to the others should cause us no concern. On telic continua, though the full conformity step may include a range of individual differences, it should not be thought of as *within itself* containing degrees; for within full conformity there are no degrees of conformity. It may vary over a greater or less range of the empirical data. The only essential fact is that all empirical units within it represent full conformity.

continuum of time: the time values are only classificatory indices for regrouping our empirical data so that they can be plotted on its successive steps. When the data are replotted, the individuals falling on the extreme left-hand step will be those who fully conform with the implied rule which expresses the accepted norm and purpose. This left-hand step gives us the terms in which the implied "rule" is set.[8] The intervals to the right of this first step embrace diminishing degrees of conformity to the rule and its purpose, until we reach the extreme right-hand step, in which there is the least recognizable degree of conformity or fulfillment of the purpose.

Let us recall again the criteria by which we decide what this last step shall be. The conformity continuum may be said to end at the last point where the individual's behavior shows that he is in some degree trying to carry out the purpose of the rule. Behavior which shows no trace of trying to do so lies beyond the end of the continuum. The individual who talks incessantly in the library, for example, from the opening to the closing of the building, shows no trace of trying to fulfill the rule of short conversations or the ends for which that rule was established. Considering the matter from the standpoint of the conformity-producing agent, we may say that *he* will regard an individual as not falling on the continuum when it is necessary, in order to make such a conversationalist pay any attention to the rule, to introduce a new element of coercion—such, for instance, as getting a library attendant to compel silence. It will be seen that such methods are equivalent to a recognition that the offender does not intend to conform at all to the purpose of avoiding the giving of annoyance and that the only way to secure compliance is to introduce the coercive act of an outside authority into the situation. The time value corresponding to this act or state of affairs is thus clearly off the continuum. These considerations require that act E in our present annoyance scale be regarded

[8] It is interesting that a procedure of this kind enables us to formulate more definitely a rule which might be otherwise merely implicit in the situation. Though no rule exists in print at the Syracuse University Library on this point, there seems to be a vague acceptance of it as a sort of "telic reality." Attendants at the library, when questioned, thought a rule did exist, and were somewhat surprised to discover their own inability to state a reference to it or to its source. The director of the library said that a rule did exist, though it was *not written*, and that it was to the effect that students should not converse in the reading room to an extent which would disturb others who wanted to work!

as beyond the end of the conformity continuum at the right. In the replotting of the data we have, therefore, indicated this fact in the graphs. The last step of annoyance-avoidance is the one just before that in which act E occurs. Cases falling in the annoyance interval in which E occurs are "complete non-conformists."

Replotting the Data Upon the Hypothetical Continuum of Conformity

The conformity continuum of the avoidance of giving annoyance is now defined in terms of its equivalent time values, and ready for use. It will be remembered that our main quest was to discover whether true conformity was present in the uniformities of observed conversation lengths in church, library, and club, and, if so, to find upon what variable such conformity lies. Reasoning from the J-curve hypothesis as an assumed basis, it was decided that, since a field of conformity (purpose, rule, 50-per-cent uniformity on the rule) and the revelant conformity continuum for distributing the data would yield a single J-curve, if we try a *hypothetical* conformity continuum and find thereby both a conformity field in the population and a single J-curve on the distribution, we could then justly infer that our hypothetical conformity continuum was the true one. We should then know not only whether conformity exists, but how great is the conformity, and in what aspect or meaning of the situation it lies.

We come now to the crucial test of replotting our data upon our hypothetical continuum of conformity and discovering, first, whether a conformity field results (50 per cent or more on the full conformity interval), and, second, whether the entire distribution takes the form of a single J-curve. If both these conditions are met, our problem will be solved.

The raw data of Tables 1, 2, and 3 have been regrouped into the time intervals corresponding to the units of the hypothetical conformity continuum. These frequencies and their proportions are presented in the last two colums under each situation in Table 6. The proportions are plotted in Figures 6, 7, and 8. An inspection of these figures will show that the hypothesis—namely, that a true conformity field and a single J-curve distribution will be revealed by the use of a conformity continuum in

TABLE 6

CONVERSATION-TIME VALUES OF THE UNITS OF THE CONFORMITY CONTINUUM OF ANNOYANCE-AVOIDANCE, TOGETHER WITH FREQUENCY DISTRIBUTIONS OF OBSERVED CONVERSATIONS

CHURCH				LIBRARY				CLUBROOM			
Units of Annoyance	Time Values in Conversation Length	No. of Cases	Per Cent	Units of Annoyance	Time Values in Conversation Length	No. of Cases	Per Cent	Units of Annoyance	Time Values in Conversation Length	No. of Cases	Per Cent
(1)	0–1′:06″	168	84.0	(1)	0–1′:15″	463	57.9	(1)	0–1′:24″	108	27.0
2	1′:07″–1′:21″	10	5.0	2	1′:16″–1′:30″	114	14.12	2	1′:25″–1′:51″	94	23.5
3	1′:22″–1′:45″	9	4.5	3	1′:31″–1′:54″	103	12.9	3	1′:52″–2′:15″	74	18.5
4	1′:46″–2′:06″	6	3.0	4	1′:55″–2′:21″	45	5.62	4	2′:16″–2′:33″	28	7.0
5	2′:07″–2′:27″	2	1.0	5	2′:22″–2′:48″	28	3.5	5	2′:34″–2′:51″	24	6.0
6	2′:28″–3′:10″	2	1.0	6	2′:49″–3′:18″	13	1.62	6	2′:52″–3′:08″	13	3.25
7	3′:11″–3′:51″	2	1.0	7	3′:19″–3′:49″	12	1.5	7	3′:09″–3′:36″	11	2.75
8	3′:52″–4′:27″	1	.5	8	3′:50″–4′:22″	10	1.35	8	3′:37″–4′:09″	11	2.75
		200		9	4′:23″–4′:53″	5	.62	9	4′:10″–4′:42″	16	4.0
				10	4′:54″–5′:26″	3	.37	10	4′:43″–5′:15″	7	1.75
				11	5′:27″–6′:10″	2	.25	11	5′:16″–5′:51″	4	1.0
				*				12	5′:52″–6′:54″	4	1.0
				12	6′:11″–7′:10″	2	.25	13	6′:55″–7′:57″	4	1.0
						800		*			
								14	7′:58″–9′:00″	2	.5
										400	

* End of the Conformity Continuum.

terms of the avoidance of annoyance—is confirmed. This confirmation, however, extends only to the church and library situations. In the clubroom no conformity field was found (first step had only 27 per cent); hence the use of a conformity con-

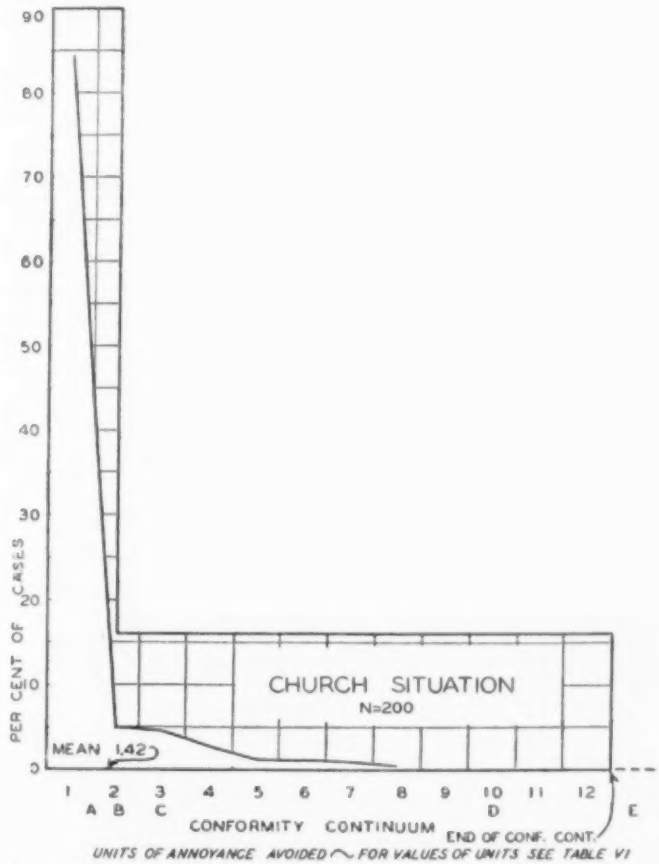

Fig. 6. Distribution of Lengths of Conversations in Church upon a Continuum of Conformity

tinuum would be irrelevant and would not be expected to yield a J-curve distribution. Nor does it yield one. We have evidence, in other words, that there is a conformity situation in the library reading room and in the church service, but not in the fraternity living room; and that the conformity in church and library lies in the fulfillment of the implied "rule" and purpose of avoiding

* End of the Conformity Continuum.

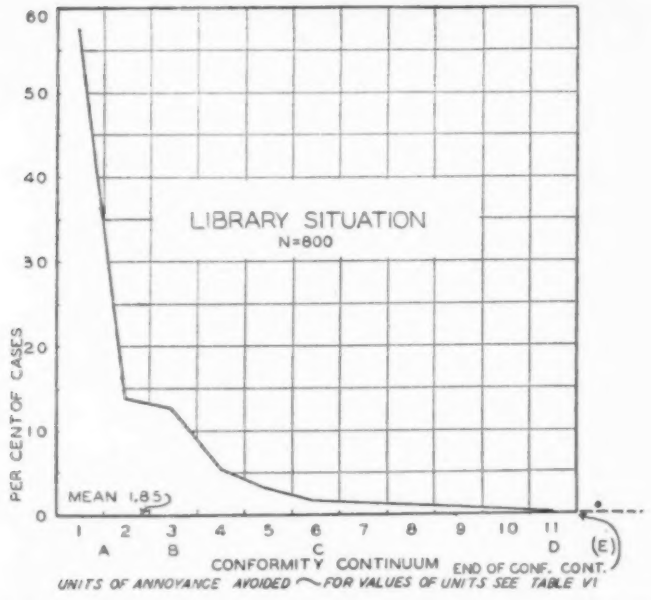

FIG. 7. DISTRIBUTIONS OF LENGTHS OF CONVERSATIONS IN A LIBRARY UPON A CONTINUUM OF CONFORMITY

annoyance to others through length of conversations. Eighty-four per cent of the cases fall upon the first step (i.e., complete conformity) in the church situation, and 57.9 per cent fall upon the first step in the library. With certain exceptions, the remainder of the proportions in both these distributions tends

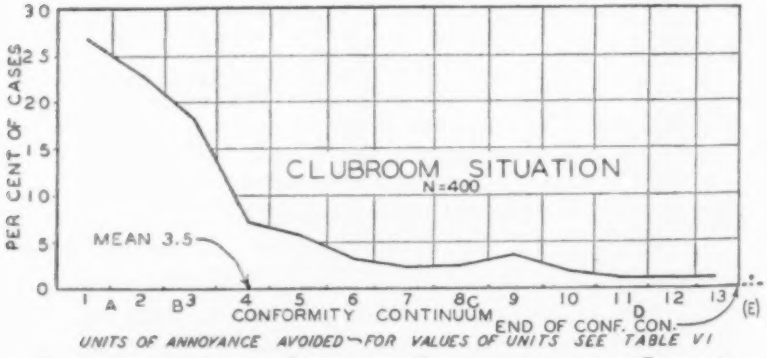

FIG. 8. DISTRIBUTION OF LENGTHS OF CONVERSATIONS IN THE CLUBROOM UPON A CONTINUUM OF CONFORMITY

to diminish on successive steps to the right of this step by increasingly smaller decrements to the end of the continuum range. Though the correspondence with the theoretical J-curve of continuous positive acceleration is not perfect, the variations are so minor that the general condition may be said to have been fulfilled. All of the population but two (.25 per cent) in the library situation fall on the conformity continuum; that is, they conform in some degree to the conformity rule and purpose. In the church situation the entire population lies well within the conformity continuum. It could, in fact, be distributed through four whole intervals farther to the right, that is, through four steps of less adequate conformity, and still be within the conformity range.[9]

To consider one of the larger exceptions to the J-distribution of the data in church and library, step 3 of the continuum has, in each case, a slightly greater proportion than that which would satisfy the demand of continuous positive acceleration. On the library curve (Figure 7) the proportions become negatively accelerated from step 3 to step 2 with relation to its preceding rise. Roughly, 30 of the 800 cases would need to be removed from step 3 to bring the curve into positive acceleration in this region. In the church (Figure 6), an increase of 4 or 5 (out of 200) upon step 2 would bring step 3 into the proper relation. These irregularities, at least that for the library, are probably not due, however, to an insufficient number of cases in the sample. Nor are they due, in all probability, to inaccuracies of the scale (ambiguity values are here low).

It must be remembered that there are other considerations which must tend to set the length of conversations, in institutionalized situations such as these, besides the avoidance of annoyance to others. Work within a library may be facilitated not

[9] It is important not to confuse this fact of being within the conformity-continuum range with what is ordinarily meant by conformity to a custom. From this latter standpoint, conformity is an all-or-none affair: one conforms either adequately and fully, or else not at all. On our conformity continuum, on the other hand, only those who fall upon its first step at the left are really "conformers" in the usually accepted sense. Our present treatment, however, is more discriminating. It accords better with the facts in taking into account those degrees of the act (on its meaning variable) which truly occur and are degrees of the purpose fulfillment, though their degree is not great enough for them to be included as within the range of performance generally accepted as adequate in the situation. Hence an act may be upon the conformity continuum (as partial or incomplete conformity), even though it is not an instance of conformity in the usual dichotomous sense.

merely by restricting talking, but, to a degree, by promoting it through occasional questions concerning assignments, references, and the like. Purposes of this sort may explain why a number of individuals talk slightly beyond the expected time in spite of the realization of giving some annoyance to others. Conversations probably do not show a great deal of elasticity with respect to time. It takes about so long to tell a thing and get an intelligible reaction. There is probably a certain uniformity growing out of the very nature of a conversational encounter between two people. It is also possible that increases in expected proportion of time of talking may be due to other continua of conformity and other smaller conformity situations operative within the larger situation. The conformity of students with the instructions of professors in getting their assignments straight might be a case in point. Where the variable chosen (here, in terms of time) is one which is also a function of other uniformities, the resulting curve must become somewhat complex. It may contain smaller modes, or even smaller J-curves, within the larger, somewhat irregular J-curve distributions. The problem of analyzing, through one empirical variable such as time, a whole pattern of interweaving conformity situations in a cross section of population at a given moment holds interesting possibilities.

Turning to Figure 8, that for the clubroom situation, we note that there can be said to be no conformity situation or "custom" with regard to lengths of conversations. Only 108 of 400 cases, or 27 per cent, fall upon the first step of our conformity continuum. The curve which results is not a single J-curve, but has negative acceleration toward the mode. There are two cases (.5 per cent) which fall off the conformity continuum at the right. The fact that nearly all the cases lie within the conformity range does not establish the conformity, since the other criteria, the modal proportion and the shape of the distribution, are not met. It is true that we do have an accumulation on the first 3 steps of the scale of 69 per cent of the cases, the curve dropping thereafter to small proportions. This accumulation, however, is probably due to other factors than shortening the time through the avoidance of annoyance to others. The natural shortness of many conversations, such as those of exchange of greetings, or questions and answers, would probably account for

some instances. If we refer back to the empirical distribution for the clubroom (Figure 3), we shall see that the slopes of the curve to the left and right of the mode are more gradual than for the other situations. This fact suggests that differences of personality, diverse interests, and circumstances of the moment may play a larger part in determining the distribution than in church or library, where like interests and conformity-producing agents are at work.

The differences in the distributions for these three situations on the conformity continuum bear out our general impressions from familiar experience. Because a conversation during religious services is likely to be more disturbing than one among students working in a library, we should expect conformity to be steeper in the church than in the library. In addition to distraction, there is the added annoyance of a suggestion of irreverence on the part of those who converse. We should expect that disapproval of long conversations in the library, however, would be more characteristic than in the clubroom, where conversations usually not only fail to disturb, but constitute one of the very purposes for which the clubroom is maintained. The means of the distributions on the conformity continuum tell the same story. That for the church is 1.42 (11.8 per cent of the scale range); that for the library is 1.85 (16.8 per cent of the scale range); and that for the clubroom is 3.5 (27 per cent of the scale range).

The conclusion that a conformity situation exists in church and library with respect to avoidance of annoyance through conversation length does not imply that there may not be other empirical variables which, if employed in collecting the data, would yield uniformities, or other conformity variables which, if used in replotting, would disclose conformities. The value of our present method, however, is that once the investigator has selected such a trial continuum, it is possible for him to put it to the test. Though other conformity continua and situations may exist within church and library conversations, the point significant for our present purpose is that we have found at least *one;* and that one alone is adequate to reveal the presence and meaning of a high degree of conformity.

INTERPRETATION OF FACTORS AFFECTING THE EMPIRICAL DISTRIBUTIONS

Let us now give brief attention to some of the influences which may have determined the approximate double-J form of the original empirical distributions. As we have noted, the distributions upon the *empirical* continuum for the church and library (Figures 1 and 2) assume, with some irregularities, the form of the double J. We may suppose that the mode of the curve, which would otherwise express unhampered personality differences of sociability, loquacity, etc., is here pushed toward the left by the real or potential influence of conformity-producing agents, or by the spontaneous desire of individuals in these situations not to give annoyance to others through talking. This common pressure probably obscures certain differences of personality which might otherwise appear. To this degree, therefore, institutional settings like library and church tend to abridge the expression of natural individual differences. The clubroom setting, on the other hand, shows a nearer approach to the normal curve of individual variations. Yet there probably are some individuals whose impulse to talk is so strong and characteristic that they will continue their conversations in spite of a pressure from conformity-producing sources. The conversations of such individuals probably make up many of the cases which extend well out on the right-side, producing the "tail" of the J. We must recall, however, that on both sides of the curve simple chance may also play some part.

When we look at the left side of the empirical distributions we see a similar phenomenon. Many did not talk as long as they could have talked and still be within the mode of conformity. Both chance circumstances and personality differences (such as traits of taciturnity and the like) may have played a part here. We have mentioned also the probable tendencies of many conversations to be short by the very nature of the case, or because of other conformity-producing influences operating in the situation. The short dashed line x in Figures 1, 2 and 3 indicates the time position where the first step of the conformity continuum (that of complete fulfillment of the purpose) ends, that is, where the *lesser* degrees of conformity begin. All on the left of line x are fully conforming to the expectancy of avoiding annoyance

through length of conversation. It will be seen that the actual mode of the distribution falls to the left of these points in library and church, but to the right in the clubroom. This fact may suggest that individual uniformities of interest in working or worshipping in themselves produce a mode of conversation time shorter than the conformity norm limit. Or it may suggest that an individual in talking is over-apprehensive of causing annoyance, as compared with the time standard used by the raters in making the scale. If this supposition is true, we should perhaps conclude that two standards of conformity may be needed, from the viewpoints of participants and conformity-producing agents, respectively. It is also possible that we should have begun laying off our units of felt annoyance a little lower down on the Y-axis (Figure 5B) than the position of act A.

Inertia, the energic factor, and other biological influences probably do not play so significant a part in the conversation-time distributions as in other conformity distributions, since the act of talking is not initiated by the conformity-producing agents, but is one which the individuals would enter upon anyway of their own accord.

Summary

A. Conclusions regarding conformity in the lengths of conversations. In the first section of this paper the prevailing methods of studying conformity, by such conceptual tools as customs, folkways, and mores, were criticised. It was contended that inaccuracies could be corrected, confusions clarified, and important questions answered, if we would consider "rule" and "custom" to be essentially multi-individual situations and would study them by observing and measuring the behavior of the individuals involved. We are now ready to sum up the case we have made, so far as our studies of conversation are concerned. Such a summary can be achieved best by restating and answering the questions raised in the first section of this article.

1. In what respect are the behaviors in the situation similar? Within what aspect does uniformity lie?

There is some uniformity, by no means complete, in that there is a limiting of conversations in church and library to a fairly short *range of time* (church, within the limits of 0 to 4 minutes

20 seconds; library, within the limits 0 to 7 minutes 10 seconds). Uniformity is thus much greater in the church than in the library. The clubroom shows some uniformity but considerably less than the church or library (range, 0 to 9 minutes). That uniformity with regard to time is present is better demonstrated by a fairly steep mode in each situation, showing a fairly large number of cases within a single 10-second interval. This mode, even for the highest (church), does not include, however, more than 27.5 per cent of the conversations. In the church, the conversations of 55 out of 200 individuals lasted from 20 to 30 seconds. In the library, 124 out of 800 persons conversed from 40 to 50 seconds. (Distribution on the *empirical continuum*.)

2. *Is there a field of conformity?*

Yes, for the church and library; no, for the clubroom. Using a *conformity continuum* based on a rule and purpose, we find that more than 50 per cent of the cases fall on the first step of this continuum for the two former situations, but less than 50 per cent in the latter; and that the two former situations yield a J-curve of distribution, whereas the latter does not.

3. *In what aspect does the conformity lie? That is, in what do the individuals conform?*

In the avoidance of giving annoyance to others through the continuing of their conversations. This is the *telic variable or purpose* which, when used as a trial continuum, revealed a situation and distribution of conformity.

4. *Are there recognizable degrees of conformity, or is there only a dichotomy of conformity and non-conformity?*

There are degrees. We should recognize two main divisions: full conformity, within which there are *no* degrees, and lesser conformity in which there may be many degrees. In addition to these there are, in one situation, individuals who show no conformity whatsoever. Conformity is thus not a dichotomy, but a *field* or *pattern*. The form of distribution of this field is that of a single J-curve (the *continuum of conformity and observed cases* falling on its different steps).

The following questions apply only to the church and library, where true fields of conformity were found.

5. *What proportion of the population fully conform?*

For the church 84 per cent, for the library 57.9 per cent (proportions on step 1 of conformity continuum).

6. *What proportion of the population conform in a lesser degree?*

For the church 16 per cent, for the library 41.8 per cent (proportion on steps to right of step 1 of conformity continuum).

7. *What proportion of the population do not conform at all?*

For the church, none; for the library 0.25 per cent (proportions falling completely off the conformity continuum).

8. *What is the relation of individual differences to the shape of the conformity distribution?*

The answer to this question is not finally known for these situations. Judging from investigations with other types of behavior, it is likely (1) that individuals possessing high degrees of talkativeness, sociability, or other traits which deviate considerably from the average, fall relatively far out on the steps of lesser conformity; and (2) that the fact of conformity means a less complete expression of natural individual differences, particularly near the mean. (*Comparison* of the *conformity distribution* with the *empirical distributions,* and of both with a hypothetical *normal curve* of individual differences. The construction and use of suitable personality continua in these situations would probably help in answering this question.)

9. *What clues are given as to probable conformity-producing agents and their motives?*

Our knowledge is here inferential. Such agents are probably (though not exclusively) the large numbers of students who read or study in the library, or members of the congregation who are participating in the service; and their motive is the desire not to be disturbed at their work or devotions. (The *logic of selecting the conformity continuum:* the choice of the continuum was based upon this assumption, and the continuum was found to be correct—i.e., it yielded a true conformity field and distribution.) Suggestions of other possible conformity-producing influences also appear (*irregularities or displacements of modes* in the conformity and empirical curves).

10. What degrees of individual variation are permitted within the limits of full conformity?

The expression of whatever degrees of the traits that affect length of talking, up to a limit of conversation length that would begin to annoy others. These limits are 0 to 1 minute 6 seconds for the church, and 0 to 1 minute 15 seconds for the library. This fact means probably that the conversational traits of the more taciturn and the "moderate" talkers only are fully expressed in those situations. (*Distributions on the empirical continue* shown to the left of the line x. Here again, a personality-trait continuum would have been an aid.)

11. What does the variable upon which the conformity is based mean in terms of the interests, motives, or behavior of the individuals who conform?

Our knowledge again is inferential. It probably means that in conforming individuals are, first, avoiding the disapproval or overt antagonism that would be directed against them by others who wish to study, by library attendants, or by worshippers in the church; second, that they are helping to make possible a general situation in which they themselves, in the capacity of readers or worshippers, may be free from similar annoyances; and third, that they are performing the natural human activity of considering the welfare and feelings of their fellows. Under, or in addition to, the last-named motive may be added the desire to maintain decorum, order, or a *milieu* of reverence. The acts through which the individuals conform, therefore, mean something to them as individuals. (Some evidence for this view is afforded by *the logic of the selection of the conformity continuum*. In its selection the motives of the participants, who in other roles are also conformity-producing agents, were taken into account.)

B. Resumé of the conformity-analysis method. It may be useful now to summarize the method which we have illustrated in the present investigation. Its successive steps are as follows:

(1) The selection of the situation and the population comprising a probable conformity situation such as we wish to study. Sociological terms denoting "customs," "folkways," and "mores" may be of service here.

(2) The observing and recording of an adequate sample of specific instances of behavior in that situation, either acts as

wholes, degrees of acts, or results of acts; these data to be classified and, if possible, measured on empirical or non-empirical continua.

(3) The selection of a variable for a hypothetical telic, or conformity, continuum by the help of such logical clues as the situation affords.

(4) The construction of a scale upon this conformity continuum variable; and, if possible, the relation of this scale, through psychophysical or other methods, to the values of the original empirical or other relevant continuum.

(5) The re-grouping of the raw data in terms of the conformity continuum intervals, and replotting of them upon this continuum.

(6) The inspection of the resulting distribution to see whether it yields both a conformity field (50 on the first step) and a J-curve of positive acceleration throughout. If these conditions are not fulfilled, a new trial is to be made using a different variable for the hypothetical conformity continuum.

(7) Assuming that a "correct" conformity continuum has been found, the distributions both on this and on the original empirical or other non-telic continua are studied and compared with a view to answering basic questions, such as the eleven here presented. For this portion of the study new exploratory continua upon various behavioral variables may be constructed.

It is to be understood that the nature of the problem and other practical circumstances may require alterations in the various steps of the method here described.

C. Concluding statement. To some readers it may seem that throughout the foregoing study we have been discovering the obvious. Everyone knows, these readers will say, that the purpose of a library is for reading and study, just as that of a church is for worship. Everyone should know in addition that since long conversations are a distraction in these places, they will be disapproved and therefore infrequent, that conformity will consist in affording the minimum of annoyance to others in this regard. To this expostulation we reply that science is frequently, of necessity, a proof and a restatement of the obvious. The assertions just made are, however, really subjective generalizations. One who makes them "feels" or "believes" that these are the conditions which *should* obtain in church and library; for they

are consistent with our awareness of the purpose of those places. A subjective sense of propriety or the rule of an "institutional authority," however, is not a demonstration of fact. It is only when we get outside this teleological view, only when we observe, and demonstrate that the things we are observing express constant relationships capable of formulations in laws, it is only then that we may lay claim to an accurate knowledge of the world about us. As we have already shown, some of the most elementary and practical questions about conformity in conversational behavior could not have been answered without an investigation such as the present. It is the hope of the writers that this method, worked out in the present relatively simple situation, may be profitably extended to the study of less tangible "customs," "traditions," "laws," and institutional situations generally, and that it may prove a first step toward a genuinely scientific study of the phenomena of social conformity.

A METHOD FOR EVALUATING THE EXPERIMENTAL USE OF DRUGS IN PSYCHOPATHOLOGY *

BY MILTON M. PARKER
The Ohio State University

The Problem

Within the last few years the application of drug therapy to the various behavior disorders has become increasingly popular. The prospect of using a specific to restore to normal the behavior of a patient gone awry is certainly something to be desired. With the exception of instances where a specific causative agent has been demonstrated, as in syphilitic infections, however, no true specific has appeared for any syndrome.

Although much has been written concerning "the individual as a whole," any experimental attempt in psychopathology rests on a specific discipline. And this discipline invariably is based on certain fundamental postulates. Consequently, while it is wholesome to speak of the unity of the human being as an organism, a method of experimental investigation must necessarily deal with only a segment. Thus, for the biochemist the problem of behavior disorder presents itself as a phase of aberration in bodily chemistry, whereas for the experimental psychologist this same chemical aberration is a behavioral phenomenon.

Between these two fundamental approaches stands the clinician with his own perennial problem, namely, "What can be accomplished in a practical way that will cure or alleviate the psychopathological symptoms of the patient?" This problem is pressing for all concerned—but for the clinician most of all. It is he to whom the layman comes as to an oracle, hoping and praying for the magic word or potion that will help; and it is the clinician who must supply this word or potion. What, then, is the answer?

Depending on the training of the clinician (psychologist or physician) or the type of approach with which the clinician has

* This paper was read before Section I of the A.A.A.S., December, 1938, and was originally presented in a seminar on "Experimental Approaches to Abnormal Behavior" conducted by Professor W. L. Valentine.

experienced most personal success, the layman receives either the word or the potion while the clinician himself may now recite the prayer. The point is that most of the experimentation which has been conducted with the use of drugs in abnormal psychological states has not yielded evidence conclusive enough to be followed with reliance.

The methods and results of experimentation have been reported with so many qualifications that anyone interested in their application would seem forced to resort to the proverbial "shot-gun" technique, purely for the sake of practicality. This problem, then, of the use of psychology or drugs or both in treatments of behavior disorder, is complex. But it is not at all insoluble; for experiments which have already been carried out, though inadequate, point the way.

Our main problem is to discover, if possible, the basic mechanisms responsible for the disordered states. If we can uncover specific therapeutic techniques which are successful with these states, we can then work back in a detailed effort to describe the original mechanisms.

CRITIQUE

The Necessity for Controls. To date, most of the experimental studies indicate that psychotherapeutic effects may be as efficacious as the administration of drugs. Other experimenters recommend that a particular drug under investigation should be used in conjunction with psychotherapy. In general, there have been no clear-cut indications as to just what the specific effect of the drug or the psychotherapy might be, nor any hint as to how each of these agents exerts its influence when used with the other.

Just as the practical use of a drug in a psychopathological condition requires two approaches, the physiological and the psychological, so the experimental problem in this field demands at least two controls. Yet in only a very few studies has such a plan been executed or even approached. It follows, therefore, that the results must often be equivocal on two counts. In patients who recover one cannot tell just what specific effect, if any, the treatment had. In most of these cases it is as correct to attribute the recovery to suggestion as to a particular physiologi-

cal effect—a result which is inevitable in the face of a lack of proper experimental controls.

The Role of Suggestion. Since the suggestive process, by its very nature, is subtle to the point of being elusive, it is not surprising to see how little the factor of suggestion has been experimentally controlled in this type of research. There is, however, good reason to assign it great importance; for while there may be some divergence of opinion, there is also considerable agreement that the suggestive process plays a fundamental role in the precipitating conflict. The reasoning is as follows: if a particular behavior disorder is viewed as resulting from a conflict, then it seems justifiable to assume that the very inability of the individual to attain an adjustment, directly opposed to the one which results in the breakdown, indicates a peculiar susceptibility to a suggestive process. With suggestion very likely playing such a major role in the initial maladjustment of the organism, it does not seem too much to expect that this same adjustment could be reordered by a use of the same process, but proceeding in the opposite direction. Witness, for instance, the success of psychoanalysis and other methods of so-called "mental hygiene."

It would be foolish to attempt to argue the supreme importance of the psychogenic nature of mental disorder, and such is not the intention here. It should be recognized, nevertheless, that a psychological process as important as suggestion must at least be accorded experimental recognition in a type of research which is, after all, dealing primarily with psychological states.

The literature in this field is so extensive, but on the whole so uniform, that nothing would be gained if more than one or two papers were critically examined. The criticisms which follow apply to almost all the papers on this topic. There are always, of course, some exceptions; but no experimental design which the writer has been able to find has considered all the important factors involved.

The Usual Sample and the Conclusions It Allows. Often the number of subjects employed in any given experiment has been too few to allow any kind of statistical interpretation. Where statistical analysis has been employed at all, by far the most widespread measure of efficiency has been based on percentage-recovery. In a type of experimentation such as this, where the

mechanism of the disorder is not yet sharply defined and where it is often impossible to discern on an *a priori* basis whether one or more mechanisms may be involved, the use of proper statistical procedure is of importance. It can facilitate the recognition of the value of the immediate data and, more important, the perception of the logical "next steps."

Wortis (**14**) has reviewed the results reported by a number of European clinics on insulin-shock therapy. In a survey conducted by the *Psychiatrisch-Neurologische Wochenschrift* (reported by Wortis) we find that of a total of 77 cases, 11 recovered, 26 improved, 38 remained unimproved, and 2 actually became worse. If then, on the basis of these totals, we draw up percentages, we see that only one in 7 patients (14 per cent) recovered, one in 3 patients (33 per cent) improved, and slightly more than one in 2 patients failed to show improvement at all. In other words, under this method of treatment the patient had, at best, a fifty-fifty chance to recover. Such results are definitely not impressive and indicate that on the basis of such experimentation we are far from realizing a specific therapy with the use of insulin. These results, however, do not necessarily discount the possible value of insulin-shock treatment. They are simply illustrative of the type of result which has been obtained by the use of drugs in general.

On the other hand, we are nearer the truth than any percentage-recovery statistics can show. But first we must cease to expect crucial results when a specific drug therapy is applied in blanket fashion. The fact that certain patients respond to a given type of therapy, while others do not, is itself indicative of difference. The method of analysis by percentage-recovery, however, cannot deal adequately with this circumstance in a manner which will show anything more than the bald fact. What is needed is a method which will, in addition to showing the proportion of those who respond to a given treatment to those who do not, test the significance of the difference. But of this, more later.

The Nature of the Criterion. It is generally agreed that the criterion for a pharmacological specific is such that the substance must manifest its action in at least 90 per cent of the cases. The only instance where any other criterion of specificity would be feasible would be that in which the validity and reliability of

our diagnostic categories were proved. It is well known, however, that there is a good deal of doubt concerning at least the reliability of the present abnormal psychological categories. And the concept of validity, from the standpoint of therapeutics, seems to have remained unnoticed—if not unknown.

The use of the terms "validity" and "reliability" may appear somewhat unfamiliar in this connection. But it should be realized that the application of most of the psychopharmacological therapies involves principles whose specific physiological mechanisms are unknown. It follows that the psychopharmacologist who would investigate the application of a particular drug therapy to various syndromes finds himself confronted with the same problem facing the psychometrist in the detection of group differences. In each case the investigator must have reasonable evidence that the groups involved may actually be considered, for the purposes of the experiment, to contain individuals who may be classified as the labels would have them appear. It may be supposed that group differences could conceivably be obtained regardless of the nature of the criterion by which the groups were originally selected. It is obvious, however, that these differences would be more pronounced and of much greater significance if the criterion used in the classification were closely related to, or the same as, the criterion involved in the experimental test. To this extent the problem of determining validity, whether it be of a "mental" test or a psychopharmacological therapy, is the same.

The investigator must also be reasonably certain that the criterion for the selection of the groups is clearly defined and uniformly employed by all who claim to effect similar separations for experimental purposes. To this extent the problem of determining reliability is identical in both instances. The basic problems, assumptions, and experimental planning of the psychopharmacologist involve techniques of measurement practically identical with those of the psychometrist. The validity and reliability of psychopharmacological therapies are, therefore, two very important phases of the general problem of abnormal psychology: the application of tests or therapies to a classified population of the psychologically abnormal.

Importance of Experimental Planning. In extremely few instances where an appreciable number of subjects has been

employed has a therapeutic measure involving the use of a single drug been successful in curing all or even a majority of a given syndrome selected for experimentation. Usually not more than half the number treated show signs of recovery, and often the percentage is much less than half. If we make due allowance for experimental error, fifty per cent efficiency (or less) should indicate at least one of several things, namely: that the original diagnosis was faulty (or a defective classification); that the psychological abnormality may involve more than one physiological mechanism; that the administration of the drug is often incidental and recovery is induced by the added attention or by suggestion given the patient; or that the improvement may be "spontaneous." Added to these considerations is the fact that the same therapeutic measure applied to the same syndrome by different investigators often yields varying percentages.

Time and Age Factors. Another factor which has gone uncontrolled is that of time. In dealing with such complex sets of variables as are involved in states of abnormal behavior it is extremely necessary that the test of a therapeutic measure, either physiological or psychological, be limited to a definite period of time. This period should be calculated on the basis of the absolute amount of time required for the specific agent to act. Beyond that time, effects should be noted with extreme caution, since the probability that irrelevant phenomena may interfere and exert influence increases tremendously. Even though follow-up study is worthwhile, usually it is not justified when all the environmental conditions affecting the patient during that period are not accorded equal recognition with the possible effects of the particular therapy involved.

Occasionally much is made of the fact that disorders of recent onset seem to yield to a particular therapy in noticeably greater proportions than do those of long duration. Also, it is not unusual to find an investigator reporting that a therapeutic method which he has used has seemed obviously more successful in young patients than in older ones. Too often the point seems to be overlooked that such a result is the usual expectation from a purely statistical standpoint, and that a similar, though perhaps smaller, recovery effect would occur even if therapy had not been administered at all.

The Specific Problem. Actually we are faced with the experi-

mental problem of differentiating between those patients who would be likely to improve and those who would be likely to remain unaffected by a particular therapeutic technique. There is no doubt that there are available methods which are successful to some degree in alleviating given psychopathological states. These methods are highly specific, and in most instances are technically more reliable than present methods of diagnosis. Until now the procedure has been to attempt to make the diagnosis and categorize the condition, then to attempt to treat it by one of the various available methods. But experimental evidence indicates that the present classification of behavior disorders is too crude for us to expect any experimental consistency from it when subjected to highly specific and reliable therapeutic methods; and that this classification, therefore, requires a revision of criteria from gross signs of overt behavior into therapeutic types defined on the basis of amenability to particular methods of treatment.

THE PROPOSED METHOD

Now how can we set up a more crucial experiment, designed to lend itself to a more fruitful analysis? The general method recommended here is the "Chi-square Test" (2; 4; 11; 13), the application of which is now illustrated.

For our example let us take a study reported by Monnier in *Nervenarzt* (8). One hundred twenty-seven schizophrenics were treated by prolonged sleep with Cloetta's mixture. The results showed that 53 patients improved, 58 remained unimproved, 14 were too weak to go through with the experiment, and 2 died. If we exclude those 16 patients who did not complete the experiment, then 53 of 111 patients, or 43 per cent, improved, and 52 per cent did not. Notice that these percentage results do not indicate any specific action.

A Hypothetical Control. Let us suppose, however, that at the beginning of the experiment another group, resembling as closely as possible the original experimental group, had been selected for observation during the experimental period and subjected to the same routine, but not actually treated. At the end of the experiment we could then compute the number of recoveries in the control group and use these data as a basis for comparison. For the sake of illustration let us assume still fur-

ther that in a control group of 125, 40 patients, or 32 per cent, improved while 85, or 68 per cent, remained unimproved. (These hypothetical figures for recovery may be somewhat high, especially when we consider the true recovery data in the group actually treated with Cloettan. But let us see what this imaginary case will show.) Comparing, we see immediately that the experimental group showed a greater improvement by 16 per cent.

Now what does this 16 per cent advantage represent? Is it a real advantage? Can it be expected to appear whenever treatment is given as above? If so, then, although 16 per cent is a small advantage, we may have a specific for at least a certain portion of schizophrenics. This portion could subsequently be studied for more valid diagnostic symptoms which, on identification, would give a basis for a true classification and specific treatment. From this point it would be possible to proceed also to a similar experimental analysis of the remainder of the schizophrenic group using other therapeutic techniques, and so on.[1]

To return to our example, we note that there is an experimental group and a control group, a group which recovers and one which does not. This circumstance then gives us four possible classes into which we group our data. Solving, we find a Chi-square value of 6.1 with an equivalent P value of approximately 1.5 in 100 that the difference between the two groups could be due to a chance selection of the particular samples. In other words, allowing for experimental error, results such as these would indicate that we probably have a specific for a certain portion of schizophrenics, notwithstanding the percentage-recovery of the actual experimental group which seemed to indicate only a fifty-fifty efficiency for this therapeutic method.

Let us remember also that the figures for the control group cited immediately above are only assumed since there was, in fact, no such group included in the original experiment. Consequently it may very well be that in an actual experiment employing the plan recommended here, much less than 32 per

[1] A precaution must be pointed out here, however. Since we do not yet have available any indication of the therapeutic validity of such categories as schizophrenia, it would undoubtedly be of more value to investigate the psychiatric specificity of a particular drug by sampling patients from all other abnormal categories as well. This would insure thoroughness and at the same time would prevent a biasing influence from entering the data as a result of sampling.

cent of the control group would show improvement. Such a result would serve only to emphasize the specific action of the drug.

An Actual Though Inadequate Control. For evidence which tends to corroborate this assumption we may refer to a study on the "Use of Histamine Phosphate and Peptone Solution in the Treatment of Neuroses and Psychoses" (Marshall and Tarwater, 7). In this study a group of 35 patients, consisting mostly (27 patients) of schizophrenic and manic depressive personalities, was treated with subcutaneous injections of graded doses of histamine phosphate. A control group of 12 patients was treated with injections of graded doses of beef peptone solution. The results were as follows: in the control group of 12, 2 patients, or 16 per cent, improved, and 10 patients, or 84 per cent, remained unimproved; in the experimental group of 35, 17 patients, or 49 per cent, improved, and 18 patients, or 51 per cent, remained unimproved. Again in this experimental group we notice the recovery percentage ranging somewhere around the fifty per-cent mark, and again the indication seems equivocal so far as specific action is concerned. Comparing the recovery percentages for the experimental and control groups, however, we find that there exists an advantage for the experimental group of 49 per cent minus 16 per cent, or 33 per cent. Now this percentage would seem to give promise for the use of the therapeutic method of histamine injection. But, even with so large an advantage, we are still not certain whether this efficiency of 33 per cent is likely to appear whenever this method is applied, or whether it happens to represent the result of a fortuitous selection of patients.

Again we may submit the data to the "Chi-square Test." When corrected for the small number of cases involved in this experiment, Chi-square has a value of 2.44, indicating a probability of about 12 in 100 that the difference of 33 per cent or greater, between the two groups, could have resulted from the fortuitous selection of the patients for the experiment. According to the ordinary criteria for significance, a probability of 12 in 100 that the observed results deviate from those which are expected because of the random sampling, is large enough to label the results as "not significant." On the other hand, such an approach would seem to offer much greater hope than that

which yields results showing, at best, an ambiguous efficiency of 50 per cent; especially since the setting up of a "null hypothesis"—i.e., that the drug has no psychiatric influence—is always feasible. We are, in other words, much more apt to realize a specific action under these conditions than if the results had been calculated only on the basis of percentage-recovery, and without a control or a hypothesis.

Experimental Assumptions and Their Consequences. If we compare the conclusions which we have drawn from the two preceding experiments, the advantages offered by experimental planning—as opposed to the wholesale application of therapies without physiologic, psychiatric or statistical hypotheses—become apparent. In Monnier's experiment (8), where we added a hypothetical control, we found that the experimental group held an advantage over the control group of only 16 per cent in percentage-recovery. Yet, despite this fact and the fact that the actual experiment itself (without the control) indicated only a 50 per cent efficiency, the "Chi-square Test" showed this advantage, even though small, to be significant. In other words, neither the percentage-recovery data nor the provision of a control group would have shown the real significance of the experimental findings. In the experiment of Marshall and Tarwater (7), where an actual control was provided, we found that the experimental group held an advantage over the control group of 33 per cent. Yet, despite this fact, the "Chi-square Test" showed this advantage, even though large, to be "not significant." And again the true value of the data had been obscured. The point cannot be emphasized too strongly that the primary consideration in the experimental use of drugs in psychopathology should be the reliability of the results which are obtained, rather than their magnitude.

How Important Is Suggestion? Thus far we have been concerned, in our experimental analysis, with the task of differentiating the psychiatric specificity of drugs as they are employed experimentally. As we have previously pointed out, however, there exists also the problem of determining the suggestive effects involved in the administration of any physical or pharmacological therapy. So far as the author is aware, there have been few studies employing this particular principle of a suggestion-control group.

Lindemann and Malamud (6), Myerson (9), Guttman and Maclay (3), and Marshall and Tarwater (7), as well as others, have accorded this principle experimental recognition, but in not quite the manner recommended here. The general method advocated by the present author, regardless of the type of therapy under consideration, involves duplication of the personnel as well as of the method in the suggestion-control group, omitting in the treatment of the latter only the active ingredient. Following the collection of the data for the group receiving the treatment without the drug, the Chi-square procedure should be applied in the same manner as for the experimental group.

In our first Chi-square illustration, we used as the control a group of patients who resembled the experimental group, but who received no therapy at all—only observation. In the present instance, therefore, we use again the data yielded by the no-therapy control, as the basis for comparison with the group receiving "treatment" minus the active ingredient. This procedure should show what can be expected, in the way of a recovery response, from a particular group of patients subjected to an apparent therapy.

If We Plan What May We Conclude? We have now used three separate groups of patients, each group resembling the other. And we have obtained two Chi-square values. The first, obtained between a treated and untreated group, should tell us the actual value of a particular drug therapy in a group so treated as compared with a totally untreated group. The second, obtained between an apparently "treated" and untreated group, should tell us the actual value of a particular type of suggestion which is inherent in any physical therapeutic method, in effecting recovery in a group given "treatment" which contains no recognized active ingredient.

Once we have obtained the data in the manner prescribed, we should be able to draw several conclusions: (1) If it should turn out that the advantages in percentage-recovery of the experimental and the so-called "treated" groups over the untreated group should approximate each other, and Chi-square be significant in both instances, then the conclusion would seem to be inevitable that the recovery of patients subjected to the given therapeutic agent is a response to the suggestion inherent in the method rather than to any specific effect of the drug itself. (2) If it

should happen that the respective advantages in percentage-recovery of the experimental and the so-called "treated" groups over the untreated group should differ considerably, and the Chi-square values be significant in both instances, then the conclusion would seem to be justified that one portion of the recoveries is due to the influence exerted by the drug, while another portion is to be attributed to the effects of suggestion. And on the basis of such a conclusion it would then be necessary to characterize, by diagnostic criteria, each of the two groups, to the end that subsequently it might be possible to make diagnoses which would, in fact, have clinical and therapeutic value in each case. (3) Further, if it should appear that the advantage in percentage-recovery of the experimental group over the no-therapy control should be relatively large, while the advantage in percentage-recovery of the apparent-therapy group over the no-therapy control should be relatively small, and the Chi-square be significant in the former instance but not significant in the latter, then the conclusion would be indicated that the recovery is a response to the specific influence of the drug, uncontaminated by any suggestive effects.

Limitations and Extensions of the Proposed Method. It should be remembered, of course, that this method of double-control, as outlined here, provides for testing the effects of suggestion which may be inherent in the administration of any form of physical therapy, as well as the efficiency of the actual therapy itself. It does not include the determination of the efficiency of suggestion as a method of psychotherapy. Nor does it include an attempt to determine the effects of interaction when a drug is used to facilitate any given form of psychotherapy. Granting that these problems are of undoubted importance, they are nevertheless of a secondary nature from the viewpoint of a systematic attack. The determination of the efficiency of psychotherapy as a method, or of the interaction of psychotherapy and drug therapy could be similarly carried out with Chi-square.[2]

Some Genetical Implications of the Proposed Method. There

[2] If at this point it seems that the present experimental analysis is attempting to control too many variables, it need only be pointed out that each of these variables does constitute an experimental problem. Any one investigation could conceivably include the determination of as many, or as few of these variables as are made possible by the method of experimentation, the research facilities of the institution, and the experimental planning of the investigator. Essentially the same method could be applied just as well to a multi-variable problem.

is another point of interest worth noting here. In practically the same manner that experiments to establish relationship between various drug therapies and the different psychopathological categories have met with disappointment, so attempts to study the inheritance of the same syndromes have met largely with failure, progress having been made no further than the establishment of the so-called "neuropathic diathesis." Even though the reason for the failure to uncover clear genetic mechanisms in these abnormal psychological conditions has been obvious for some time, no substantial progress has been made. "How," say the geneticists, "can we attempt to determine a genetic mechanism for a character which has not been adequately defined for us?" Such genetic attempts as have been made have succeeded in indicating one thing mainly—that the present psychiatric categories are not valid enough to yield any consistent results. If the plan of experimentation and analysis described here succeeds at all, then the system of classification resulting from it should have not only clinical and therapeutic value, but genetic value as well.

Is a Control Always Feasible? The question may be asked, "How can this method be applied in experiments where the therapy involves the production of so unmistakable a physiological effect as to make virtually impossible a simulation of the same effect in a control group?" The induction of prolonged sleep is a case in point. Either of two methods could be used to effect controls. One very obvious method consists of using two different agents to produce essentially the same effect and then comparing the results of each with the other according to the Chi-square technique.[3] Another method consists of using the experimental group as its own control. In this case the group is subjected to periods of drug therapy alternating with periods of rest. Under these conditions it is then possible to check the behavior of the group just following the therapy against the behavior of the same group at given intervals during the rest period. Thus the group could be compared with itself, according to Chi-square, under the two separate conditions.[4]

[3] Perhaps a better analysis in this instance might be obtained with the use of the "Method of Paired Comparisons." In any case this particular control, while yielding valuable information, would indicate only the relative efficiency of the two therapies. However, if two therapies produce essentially the same subjective effects but vary significantly in efficiency, is the control not performing the desired task?

[4] In the light of what has already been said in connection with the time factor and "spontaneous" remission, the latter procedure would be useful even in conditions where another control is available.

Some Psychiatric Implications of the Proposed Method. It is hoped that the method recommended here will be employed in the experimental use of drugs as well as in other therapeutic techniques in psychopathology to test the significance of results, and to help identify basic mechanisms. These mechanisms, once discovered, should then provide the groundwork for an experimentally verifiable system of classification. The reasoning is thus: if we can establish the fact that a certain proportion of abnormal patients invariably responds to a particular therapy, then it should be possible to select diagnostic criteria which will enable us to distinguish this group originally. The same method might be followed with any therapy showing a consistent efficacy, regardless of the latter's size with respect to the proportion of patients benefitted. The intention of the author is that in this scheme the new categories would involve truly operational bases, since each category would constitute a general picture calculated to yield to a specific therapeutic technique.

Is the Proposed Method Always Necessary? It should be remarked that the use of Benzedrine in the depressive disorders, especially in conjunction with psychotherapy, has been shown to be more obviously successful than most methods of treatment. These results have been especially noteworthy in studies where subjects have been selected for experimentation on the basis of generally depressed behavior rather than on the experimentally outmoded, traditional system of classification. In a well-controlled study of this type, Dub and Lurie (1) have demonstrated the psychiatric action of Benzedrine (amphetamine) so clearly, in depressive patients drawn from five different psychiatric syndromes, that application of Chi-square is unnecessary. It is needless to say that such results are an exception to the usual rule. The point worth noting is that Benzedrine seems to offer a therapeutic rationale which, in its very efficiency, cuts across the old categories. If the results of all the therapeutic methods used in psychiatry were equally as decisive in their indications, the necessity for resorting to statistical procedure, in order to theorize about the probable constitution of our psychologically abnormal population, would be greatly diminished. Unfortunately, however, this does not happen to be the case. On the other hand, it should not be reasoned that because so many of the experi-

mental therapies have been inconclusive, they therefore have nothing to offer.

Similarly applied, the Chi-square method can be made to test the efficiency of many other psychiatric therapeutic techniques. In itself, of course, the "Chi-square Test" will not solve all of the therapeutic problems of psychiatry. But Chi-square is only one of several possible approaches. The purpose of the present analysis has been to point out a system for designing experiments in one particular phase of psychiatric therapeutics. The object of this experimental system would be the delineation of syndromes according to their causal bases, as well as the provision of the usual tests for determining the efficiency of the various supposed therapies. Each of these problems, in fact, is so intimately linked with the other that it is difficult to see how one could be adequately handled without the other. The whole problem of the establishment of effective therapeutic methods in psychopathology seems to suffer from a too strong acceptance of the dogma of traditional methods of classification. In other words, since the usual syndromes are implicitly assumed to be valid, the burden of proof is thereby thrown upon the particular therapeutic methods under investigation. Is it even necessary to point out that the chances for establishing correlation between therapeutic measures and psychiatric categories are slight when the therapeutic reliability of the syndromes is uncertain and the validity of the various specific therapies unknown?

It is obviously impossible to outline in the course of one paper the exact direction which research, using the approach suggested here, should take in every instance—impossible and perhaps undesirable. The attempt has been made simply to sketch a general method of attack. It remains for those actively engaged in investigating the worth of the various therapies in psychopathology to evaluate, to modify if necessary, and to test the plan set forth here. It is the experimental test alone which finally shows the way.

We are on the threshold of an era which is bound to throw more light on the specific nature of behavior abnormality than has ever been done before. An abundance of technical procedure is available. Yet, despite this, the task is not an easy one. For, there is more involved than the application of a single therapeutic technique to a single abnormality. There exists a confusion

among the therapies as well as among the various psychopathological states. Basically the problem consists of cutting through these confusions in a systematic manner such that the results, when measured, will be precise and *definitive*. There is a cosmos in this seeming chaos, but blanket methods of investigation can never find it. Our hope lies in planned, rather than routine, experimentation.

BIBLIOGRAPHY

1. DUB, L. A., AND LURIE, L. A. Use of benzedrine in the depressed phase of the psychotic state. *Ohio State Med. J.*, 1939, **35**, 39–45.
2. GOULDEN, C. H. *Methods of statistical analysis.* Minneapolis: Burgess Pub. Co., 1937.
3. GUTTMAN, E., AND MACLAY, W. S. Mescaline and depersonalization. *J. Neurol. Psychopath.*, 1936, **16**, 193–212.
4. FISHER, R. A. *Statistical methods for research workers.* Edinburgh: Oliver & Boyd, 1935.
5. JUNG, F. T. Centripetal drift: a fallacy in the evaluation of therapeutic results. *Science*, 1938, NS **87**, 461–462.
6. LINDEMANN, E., AND MALAMUD, W. Experimental analysis of psychopathological effects of intoxicating drugs. *Amer. J. Psychiat.*, 1934, **13**, 853–881.
7. MARSHALL, W., AND TARWATER, J. S. Use of histamine phosphate and peptone solution in the treatment of neuroses and psychoses. *J. nerv. ment. Dis.*, 1938, **88**, 36–41.
8. MONNIER, M. Treatment of schizophrenia with narcosis produced by Cloetta's method. *Nervenarzt*, 1936, **9**, 14–29.
9. MYERSON, A. Effect of benzedrine sulfate on mood and fatigue in normal and in neurotic persons. *Arch. Neurol. Psychiat., Chicago*, 1936, **36**, 816–822.
10. NEUSTATTER, W. L. Some methods and problems of psychotherapy. *J. ment. Sci.*, 1936, **82**, 47–62.
11. PEARL, R. *Introduction to medical biometry and statistics.* Philadelphia: Saunders, 1927.
12. RICHARDSON, R. B., AND WEISENBERG, T. H. Psychotherapy as substitute for narcotics. *J. Amer. med. Ass.*, 1931, **96**, 1574–1576.
13. SNEDECOR, G. W. *Statistical methods applied to experiments in agriculture and biology.* Ames, Iowa: Collegiate Press, 1938.
14. WORTIS, J. Sakel's hypoglycemic insulin treatment of psychoses. History and present status. *J. nerv. ment. Dis.*, 1937, **85**, 581–595.

PSYCHOLOGICAL STATES RESULTING FROM PARATHYROID DEFICIENCY *

BY CAROL FINLEY DDA
Los Angeles City College

THERE is no field of human behavior of more interest to the psychologist than that due to the action, both normal and abnormal, of the endocrine glands in the body. Among the lesser known of these are the parathyroid glands, for they were not discovered until 1880, and diseases of these tiny glands are rare.

The most familiar parathyroid disease is a deficiency following the injury or removal of one or more of these glands at the time of thyroid surgery. The fact that a deficiency reduced the calcium in the blood and increased the phosphorus was not known until 1908. Today a diagnosis of a parathyroid deficiency rests primarily on the measurement of the calcium and phosphorus in the blood. The normal amount of calcium is usually given as 9–11 milligrams per 100 c.c. of blood, though some authorities place it higher.

Though post-surgical hypoparathyroidism is the most common type of parathyroid deficiency, it is none the less rare, for surgeons carefully avoid removing, if possible, the posterior portion of the thyroid gland in which the four parathyroids usually lie embedded. Yet though the percentage of cases of parathyroid disturbances is decreasing, the actual number of cases is increasing, because of the larger number of thyroidectomies performed. Parathyroid deficiencies will probably continue to occur after surgery, for the tiny glands are sometimes embedded so far within the thyroid that they cannot be seen, and are sometimes in diseased tissue which must be removed.

In case a parathyroid deficiency results from surgery, the reduction in the blood calcium causes a severe cramping of the

* The writer is indebted to the members of the Department of Psychology of the Los Angeles City College for their constructive criticisms and suggestions in arranging this material; and to Dr. Scott S. McCune, without whose medical care the article could never have been written at all.

muscles, called tetany, which may occur in any muscle in the body, but which tends to disappear in time, even when the disease is untreated, and to remain latent. The most characteristic tetany is in the hand, with the thumb in the palm, the fingers straightened out toward the wrist, the wrist itself sometimes bent in. This tetany will sometimes reappear in chronic cases if a tourniquet is tied around the upper arm, a test called the Trousseau sign. In chronic cases, also, a jumping of the muscles of the upper lip and nose may occur if the trunk of the facial nerve is tapped on the side of the face just below the zygomatic process. This is the Chvostek sign. The disease is known to affect the peripheral nervous system, and to cause a tingling sensation on the skin. The fact has been noted that at the onset of the disease, the patient may or may not mention some of the symptoms, but that if the tetany occurs the tingling sensation and an extreme cold are probably already present. Disorientation as to time and surroundings has been noticed in several cases immediately following surgery.

In 1909 the gland hormone was isolated, and by 1924 perfected. It was at first supposed that the hormone, when used intramuscularly, would take the place of the glands which had been removed, for its use brought the blood calcium up to the normal level; but time has shown that it does so by utilizing the calcium in the bones, and, if used alone over an extended period of time, may leave the bones soft.[1] After a time, moreover, the dose must be increased, and, if discontinued, the blood calcium is likely to fall below its former level.[2] Whether this result is due to a building up of "anti-hormones" in the body,[3] or to a degeneration of the remaining parathyroid tissue, as is the case with dogs on whom experiments have been performed,[4] is not definitely known.

Since the failure of this hormone to accomplish the hoped-for results, other treatments have been used. At the Johns Hopkins Hospital, Dr. Harvey Stone has had some success in transplanta-

[1] The advisability of using parathyroid hormone in cases of parathyroid deficiency is still controversial, but with the discovery of other more simple treatments for the disease, its use is declining.

[2] Freyberg, R. H., Grant, R. L., and Robb, M. A. Hypoparathyroidism; the treatment of chronic cases. *J. Amer. Med. Assoc.*, 1936, **107**, 1769.

[3] *Ibid.*, 1770.

[4] Shelling, D. H. *The parathyroids in health and in disease.* St. Louis: C. V Mosby, 1935. P. 108.

tion of parathyroid tissue by using a new technique; at the Mayo Clinic, massive, frequent doses of calcium and cod liver oil, by mouth, have been given with good success; and only recently, in Saint Louis, Dr. Cyril MacBryde has been successful in treating chronic cases of hypoparathyroidism with a new form of vitamin D called dihydrotachysterol and small doses of calcium, both by mouth. All of these methods seem to hold out hope that this stubborn disease will ultimately be conquered.

In medical literature reference to the behavior of patients with chronic hypoparathyroidism is infrequent and indefinite. A stumbling gait, a timidity, and a rigidity of the face (called Parkinson's mask) have been noticed. Some patients have complained of noises in the head [5] or a roaring in the ears.[6] The disease is, however, conceded to be not mental, any sign of such symptoms being ascribed to a possible hypothyroidism.

The lack of information on the psychological aspects of the disease is due in part to the rarity of the disease; it is due also to the fact that the patient himself, being timid, is afraid to tell the doctor all that he experiences. It is with the idea of telling what goes on behind such timidity that this article is written, introspective though it must, to a large extent, be. Though it is primarily a personal history, the writer was fortunately able to find another similar case, whose account is also given. It would seem that the striking similarities are certainly more than mere coincidence.

CASE 1

The following objective view of Case 1 is written by Dr. Scott S. McCune, M.D., of Los Angeles, who has studied it over a long period of time.

This patient first came under my observation on January 6, 1936, sixteen months following a sub-total thyroidectomy for Graves disease. At this time the patient was complaining of muscular cramps, and nervousness. Examination revealed positive Chvostek and Trousseau signs. Fundoscopic examination by an oculist revealed beginning lenticular opacities, bilaterally. Blood chemistry studies revealed the serum calcium levels to be significantly depressed, reaching levels as low as 6.6 milligrams per cent.

Temporary relief of subjective symptoms and objective signs could be obtained with intravenous calcium. Parathormone proved of only temporary value and the patient gradually became refractive to this hormone.

[5] Boothby, W. M. Parathyroid insufficiency. *Proc. Staff Meet. Mayo Clin.*, 1935, 10, 91.

[6] McBryde, C. M. The treatment of parathyroid tetany with dihydrotachysterol. *J. Amer. Med. Assoc.*, 1938, 3, 306.

Satisfactory control of the symptoms and signs was obtained only when this patient was placed on frequent oral ingestion of calcium lactate supplemented by cod liver oil. Under this regime the serum calcium levels attained normal limits. The objective signs disappeared. Emotionally the patient appeared to be definitely more stable. The psychological and emotional changes can best be described by the patient herself and I can only state that our objective observations seemed to be consistent with the changes described by the patient.

One of the most amazing features of this case has been the almost complete disappearance of the lenticular opacities, which is confirmed by the competent oculist who made the original observations.

In summary then, this patient presents a case of surgically induced hypoparathyroidism, which has been satisfactorily controlled by the oral administration of adequate amounts of calcium and vitamin D. This has been largely possible because of the intelligent cooperation of the patient. The emotional changes developing during this episode, I have found of particular interest and would commend Miss Dda's account of them to all students interested in the subject.

The following data, taken from the hospital chart, will help to clarify the story which follows:

Surgery was performed on September 12, 1934. No parathyroid glands were found in the specimen. Tetany appeared four days later, on September 16. The blood calcium was 6.4. At this period, according to the chart, the patient made no complaints.

In January, 1936, the patient was again admitted to the hospital. She appeared normal mentally and emotionally. After an intravenous calcium had been given, she became almost hysterical and quite hilarious, and when the objective signs of the disease again became apparent, appeared to be very depressed. After some weeks of unsuccessful treatment, the patient became emotionally uncontrolled, and very nervous. Emotional outbursts were frequent, as was hysteria. The lowest blood calcium during this period was 6.6. The final treatment consisted of one teaspoon of powdered calcium lactate, in solution, every two hours, eight times daily, in conjunction with two teaspoons of cod liver oil, three times daily. In four days the blood calcium was 9.8.

In November, 1937, when a second thyroid surgery was performed, one parathyroid gland was found in the specimen. How many parathyroids are left is unknown. The thyroid gland has been entirely removed and one grain of desiccated thyroid daily, substituted.

In March, 1939, the blood calcium was 10.9. The dosage of calcium lactate is now four teaspoons daily.

MY OWN STORY

I awoke from the anaesthetic feeling relaxed and completely tranquil. I felt no pain. I was conscious of being moved into a private room, of having oxygen given to me, of a nurse coming and going. Then I slept.

The next few days were bewildering and confusing. Sometimes hours were telescoped into brief moments; sometimes

minutes became long hours. Time had no regularity and no balance. It crept or it raced, as it chose. I remember one visit of the doctor when I closed my eyes and opened them again only to find that while my eyes were closed, his smooth-shaven face had grown a stubble. Sometimes I felt hysterical, though I did not always give way, but always clutched at the bedclothes and swallowed convulsively.

When I slept I had bad dreams. In one dream, I was ground up into pieces and mixed with cement to make a sidewalk. As the cement hardened and was walked on, sharp pains shot through me. When I awoke from this nightmare, I was drenched with perspiration and found the pains real. Sharp needles by the thousands seemed to dance over my legs, on, and just under, the skin. They played over me like rain on the roof, staccato, quick, inconsistent. After a time they became monotonous. They climbed slowly up my body, extending from the legs, over my arms, face, scalp. When I pressed the skin, the tingling would stop in that spot, only to increase again when the pressure was removed. After several hours it seemed unendurable, and I scratched and turned from side to side to escape it. I do not remember having mentioned this tingling pain to the doctor. I thought that he already knew it.

Soon, a cold numbness began to creep into my bones and, engulfing my feet and legs, seemed to afford me some protection from the pricking needles. While I felt the needles less sharply, I also felt the covers less. At times I could not be sure that the covers were really over my feet, because my feet were so numb. The hot water bottle seemed only tepid. Soon, from my waist down, I had almost ceased to feel anything but the terrible cold and the pricking of the needles. Only by moving my legs, or by looking at them, could I tell where they were. From the waist up, the perspiration rolled off me. The weather was hot, my temperature was 103; still as I sat propped up in bed, I wiped away the perspiration with cold, numb hands.

To the tingling sensation and the cold numbness there presently was added another factor. My legs and toes, beginning to cramp, soon seemed to be doubled up. Then my fingers and arms cramped. Although these cramps were tiring—for it seemed as if invisible rubber bands pulled the muscles taut—they were not painful. If I made the effort, I could stretch out

my hand and extend the fingers, but they would always go back to the same grotesque position. The muscles would not relax. The cramps extended to every part of my body; they made it difficult for me to focus my eyes, to drink or to eat. They pulled my face into a peculiar grimace, and my tongue at times cramped so badly that I could not talk.

The first night after this cramping appeared, when the light was turned out and I could no longer see the objects in the room, I found that I could not tell where I was. I could scarcely feel the bed or the covers. Added to this experience was a feeling of being suspended in midair; my body seemed to have no relation in space to anything around it. At one time I would seem to be upright, at other times upside down or on my side. When, by clutching the bedclothes firmly, if somewhat clumsily, I could make out the fact that the bed was there, it became part of me, and the bed, too, lost its moorings. I had no particular feeling of motion; the bed and I were merely first in one position, then suddenly, without moving, in another position. When I felt the bed tipped sideways, I rolled in the opposite direction and clung to the sideboard to keep from falling. Finally I rang the bell for the nurse, to ask her to leave a dim light on. "I don't know where I am, in the dark," I explained. With the dim light turned on, I felt quite secure, for when the bed started turning topsy-turvy again, I had only to open my eyes to see that everything was all right.

The next morning they came to take some blood from my arm. I held out my arm as best I could, and turned my head away. As the needle went in, I cringed. The sensation was more painful than I remembered its ever having been before surgery. And as the blood was drawn out, I felt an unaccountable desire to cry. As I sobbed, I begged the technician to stop cutting my arm. He protested that he was not doing so, urging me to look if I did not believe him. I looked. He was only doing what had been done often before, but still the arm continued to hurt. Pains shot up and down it, and even after the needle had been taken out, it continued to throb and burn and bleed profusely for some time.

Later in the day I was given an injection of calcium in the vein. The flow of calcium into the vein provoked the same emotional reaction as had the removal of blood. Even though

this time there was no pain, I felt impelled to sob bitterly. As the injection was given, the ice in my bones began to thaw out; my feet began to get warm. By the time the doctor took the needle out I was warm as toast, and relaxed. The cramps were gone. Later, when I closed my eyes, I could tell I was lying in bed; I could feel anything I touched, and, for the first time in several days, I could feel the covers over my feet. But one thing which I had noticed before still remained. I saw things around me with a clarity I had never previously known. Though I had always been easily hurt, I was sensitive now not only in this sense, but also in the sense of seeing fine shades of feeling in other people. I had known that the nurses thought me a little out of my mind; I had known that the doctors had not expected me to live. Whenever anyone came into the room, I knew almost at once what they were thinking about. And as the calcium fell again below normal and I again suffered the cold numbness, the tingling and the cramping, as I again spent dark hours insecurely balanced in space, this quality of acute awareness of what went on around me remained strikingly present.

One day, when I was feeling fairly well, I climbed from under the covers and inched my way to the foot of the bed, where by clumsy contortions on my part I could see myself in the nurse's mirror. I gasped. A more repulsive-looking face I had never seen. It was puffy and swollen, covered with hard pimples, and of a color somewhere between a sick green and a bilious yellow. It was flesh-colored with all the red shades removed. My hair, which had been soft and glossy and curly, was as dry and brittle and straight as straw. The sight shocked me. I climbed back in bed and made a complete survey of myself. My skin, dry and puffy from head to toe, was all the same indescribable color. Pimples covered me from the waist up. As I pulled the covers up over me, I felt disheartened indeed. The doctors had assured me before surgery that I was going to be very much better, but now I decided that they had not known just what they were talking about.

As days went by, I began to realize that none of the doctors knew just what to do for me. I decided, therefore, to go home upstate with my father, who had come to get me, and put myself

under a doctor there; and I insisted on leaving the hospital, even though I could barely walk.

I rested at home for several months. Four months after the surgery, I took stock of myself. I was able to walk fairly well, though my balance was poor and my sense of touch very little better. Often I could not tell where I was walking or what I was touching unless I looked. So accustomed had I become to the irritating tingling that, since it had diminished, I scarcely noticed it. The cramps had also diminished, being now mostly in the deep muscles; unless I moved too rapidly, they were not troublesome, and the severe ones were always in my feet. Because of a feeling of insecurity in the dark, I still left a light on at night. Some few pimples, which never came to a head but remained hard under the skin, still occasionally appeared. I was always cold and numb. New hair was coming in, but it too was straight and dry. Sex desire had returned with a normal monthly cycle. I was acutely conscious of pain from any cut or bruise. If I cut myself even slightly, the wound would throb and burn and often cause shooting pains.

I could reason logically, but my mental pace was very much slowed down. I did not think of myself as being nervous, nor do I believe that I was, at this stage, though at times when the slow tempo of my days was momentarily increased I became irritable. Generally, however, geared to a very slow speed, I managed with effort to appear both balanced and calm. But I was becoming more and more conscious of the mask which my face had assumed. I no longer exhibited the slightest sign of personality. I never laughed; rarely even was I able to smile. I realized that this condition was due to the tightness of the facial muscles and that it would probably remain until I was cured. There was no way in which I could express, socially, any emotion whatever, for my mind functioned so slowly that I never could think in time of the words I wished to use. I realized also that my slow pace mentally, in combination with the lack of facial expression, often made me appear stupid, and that the fact that I was much more easily hurt than ever before did not show in the least.

In the town where my father lived I had not been able to find any doctor who would do anything for me. I decided, there-

fore, to return to the city where I had formerly lived, go to work, and go to the doctor who had advised the surgery.

As soon as I returned to an active normal life, I was faced with complications of which I had not dreamed. The actual physical activity resulted in both an increased cramping in all the muscles, particularly in hands, feet and eyes, and in a greater amount of the tingling in my feet. The sense of touch remained poor, as did the sense of balance; an increase in tempo would often cause me to stumble and fall without even being conscious of the fact until I found myself headlong on the sidewalk. So much did these tumbles embarrass me that, in order to avoid them, I watched the path ahead of me very carefully when I walked. Even this precaution was not always successful, for with eyes unable to focus properly I very often had difficulty in seeing.

I found, too, that as I had to think faster, my mind often became confused. Easily discouraged, I often kept on with a plan simply because of an inability to adjust to a new situation. The attempt to make adjustments not only made me nervous, but even in minor things was a slow process. I found myself trying to plan ahead just what I would do in any one of a number of situations. I tried to foresee and be ready for the unexpected.

Shortly after going to work I became conscious of another complication. I was not then, nor am I now, aware of its beginning. It seemed to grow out of the feeling of confusion which I had had on rare occasions at home. Now, whenever I was spoken to unexpectedly and sharply, in a way which required me to defend some action of mine—a rhetorical question such as, "How can you be so silly?" very often served—I felt very helpless and insecure, with a feeling in the pit of my stomach as of everything turning upside down. Immediately a whirring vibration which seemed to be in my head would start. This vibration, so slow that I could distinguish between the beats, was yet too rapid to count. Since it reminded me of the vibration of an egg-beater, I called it in my own mind by that name. When these periods came, my whole attention was focused on myself with such intensity that I could hear only vaguely the sounds going on around me. The vibration never seemed to last more than two or three seconds, but while it lasted, all rea-

soning power was gone. I became stoical about these periods, finally; they never seemed to get any worse. As each one began, the thought would flash through my mind, "There goes the egg-beater again," and I would concentrate on it until it was gone. After each one, however, I lived in greater dread of the next one. I could see no way of avoiding these spells, which were always followed by a cramp in some part of my body.

Such periods had the effect of making me more alert, and the increased alertness in turn resulted in an almost phenomenal ability to record in my mind, as on a motion picture film, the words, gestures, and facial expressions of a roomful of people. The quality which I had already noticed—that ability to read other people's emotions and know what they were thinking—now returned in greater degree than before. Though I might prove incapable of reacting immediately to a situation, all the minute details of that situation were none the less printed indelibly on my brain, particularly if it concerned me in any way. I could remember whole conversations word for word and run the whole scene over again in my mind, at will, as one would run a talking picture.

I had been several times to see Dr. X, and, when he offered no suggestions, went to others; but no one did anything to help me. I returned to Dr. X eight months after surgery, hoping to tell him of this—to me—alarming vibration. Knowing him to be a fine doctor and a kind man, I was sure that if I could make him understand the trouble I was in, he would do what he could to help me. I was afraid that I was insane. As soon as I saw him, I found that I was afraid to speak openly for fear that he would think the same; when he told me that there was nothing the matter with me but "nerves," I thought that was what he meant. I never went to him for help again on my own initiative.

I was afraid, now, to go to any doctor for fear of being sent to an asylum. My mind already filled with vague fears, I suffered from a feeling of utter frustration which no effort of mine could conquer. I had the feeling of being in a cage. I had long since lost the ability to talk of my emotions; I now lost also, to a large extent, the ability to understand myself. Wishing to avoid meeting strangers, I went nowhere except when necessary. I became deeply depressed. I could stand the cramps,

what there were; I could stand having every little cut or bruise hurt intensely; I could stand the stumbling and the physical insecurity, the inability to see well. I could stand all the physical troubles which this disease brought, and think them nothing, but it seemed to me that I could not endure forever the emotional sensitivity which left me open for hurt at every point, nor the inability to express how I felt, nor the mental exhaustion which overcame me at times. Sometimes people would tell me that they thought me extremely unemotional and cold or heartless. Such a reproach wounded me deeply, for though I knew it was not true, I could find no words to deny it. Most of all I was in terror of these vibrations.

But I could see no way out. If I had considered suicide, which I never did, I would have had too little initiative to carry out such a plan. Many times as I lay in bed at night I tried to cry out my heartache on the pillow. I could not even cry. So heartsick and discouraged did I become that my stubborn pride alone kept me from giving up. I went on working.

Physically I continued to improve, for, finding that I was less depressed when I was less tired, I rested a good deal, and in my spare time I read widely. I studied such authors as Raymond Pearl and H. S. Jennings. Because my mind moved slowly and cumbersomely, I found it difficult at first to focus my attention, but once started I became absorbed and grasped the material better than I had ever done. As a result of all this reading—as I thought—my vision became so poor that sixteen months after surgery I saw an oculist about it. He sent me to Dr. X again, telling me that I had cataracts from hypoparathyroidism. I was happy for the first time in months. Here was something tangible which I could mention to a doctor without being thought insane. For the first time in months I had hope again. Within a few days I was back in the hospital, and ten weeks later walked out adequately treated.

The only time the "egg-beater" ever returned was once when parathyroid extract was being used. It was then definitely a sound, a buzzing, and the worst one I have ever had. It has never occurred again. The extract also made me less conscientious; in fact, until it was used I did not realize that ever since surgery I had been over-conscientious; and too, it made me insen-

sitive to the suffering of those around me, and temporarily hypocritical in my attitude of sympathy.

The final treatment has already been outlined. My initiative returned at once. For a while the physical troubles which I had had returned when the blood calcium was low—which was in the early morning; but in the daytime the cramps were gone, the sense of touch, sense of pain and the sense of balance were normal, and I was not cold. My mental reactions, though more normal, remained somewhat slower than they had been before surgery, when, of course, the speed may have been abnormal. Once more I was able to put my feelings into words in my own mind and, before many weeks had passed, to talk of them and to express emotions in both words and facial expressions. I found, too, that I had my "personality" back and could begin to make friends again; I wanted to meet people. Since my appearance was improved, with the remaining pimples and puffiness disappearing and the skin becoming normal in color, I was no longer embarrassed by meeting anyone.

The only symptoms which did not disappear almost at once were the extreme emotional sensitivity and its accompanying features: the overconsciousness and the dogged perserverance which had kept me going for so long. Over a period of two years, however, they gradually became normal. For a year after the beginning of the treatment I was again emotionally an adolescent, frequently told by my friends that I acted like a child.

More than two years after the treatment had been started, it was discontinued, and the blood calcium measured. At this time the tingling sensation, the cramps, the lack of facial expression and the blurred vision all returned to a minor degree. The vague fears and the depressed feeling reappeared, but the vibration did not. Touch, pain, and balance remained nearly normal. The blood calcium now dropped only to 8.5, which is nearly normal. The doctors said, "The remaining glands are compensating." The medication was re-ordered in smaller amounts. Today I am normal, the cataracts have nearly disappeared, and I am well.

To me, hypoparathyroidism is an extreme dissociation of the "inward" and the "outward" personalities in which there can be no harmony, no possibility of being understood and no cessation of conflict as long as the disease remains untreated.

CASE 2

Mrs. S, aged 37, appeared when interviewed to be quite normal mentally and emotionally, seemed intelligent, and answered questions readily. Most of the information she volunteered when she was informed of the purpose of the interview. Surgery had been undergone two years previously, with the removal of one parathyroid and injury to another. The hypoparathyroidism had not appeared until three weeks after the thyroid surgery, at which time the tetany had been extremely painful —"worse than labor pains," she said. No treatment has been given for the parathyroid deficiency, but the tetany is no longer very painful; and she stated that she had improved physically following x-ray treatments for a fibrous uterus. At times, however, the tetany is worse, so that she is unable to do even light housework.

Mrs. S moved very slowly, though not awkwardly, and at times seemed to have difficulty in enunciation. She seemed to understand her own condition very well. Several times she spoke of being "unable to find herself," and of the difficulty of finding words to express what she wished to say, or to express what the disease was like. "It is not mental, but it is more than tetany and things like that. It is a crazy disease. It doesn't make sense. I couldn't tell these things to a doctor. He'd think that I was imagining things or that I was neurotic, but you and I know that they are the truth. If doctors understood how terrible the disease is, they would work night and day to cure people of it, but they just don't understand."

The following symptoms are those which appeared at the time the disease began and which have continued since. Others, which she ascribed to the menopause (beginning five years ago), have been omitted.

"I have had cramps in every part of my body. . . . I was bloated and my skin was awfully yellow with a gray-green cast. It isn't all gone yet, and I hate to look at myself in a mirror. I still have a few pimples. They are hard, and seem to fester, but they never come to a head. . . . My joints are stiff and it is hard for me to move. It is hard for me to keep my balance. I lose control of my legs, and my elbows and knees knock. I cannot walk up or down stairs without holding on. I feel some-

times as if I were falling backward, and I get dizzy when I stand up quickly. I fall down in the street sometimes. I don't faint, I just fall, and don't know it till I find myself on the sidewalk. I don't sleep well, and when I am in bed at night I feel as if I were falling. Sometimes I feel as if I were upside down. . . . My hands and feet are always cold, like ice, and even the hottest hot water bottle feels scarcely warm. I am numb most of the time. I can feel things when I touch them, but they seem to be a long way off. . . .

"I had the tingling sensation for a long time—still do, a little. It was so bad that I could hardly stand water running over my hands. I still don't like to wash them. Sometimes, I seem to feel things walking up and down my back. . . . I am never hungry. I can't taste things very well, nor smell very well. At first I couldn't taste nor smell at all. I lost my voice at first, and even when it came back I couldn't carry a tune any more. I still can't. . . . When I cut myself, the cut seems to be deeper than it is—it hurts deeper, the hurt lasts longer than it used to, and always burns so much. . . . I have a floating feeling in the top of my head.

"I have a vibration sometimes. I don't know just where it is. It seems to be in my head and ears and all over my body. It's a rocking vibration, like an earthquake. I call it a pulsation. . . . I have a rapid pulse all the time, and it seems as if my heart had to push hard to do its work. . . .

"I can think all right if I don't get excited; then my thoughts get all jumbled and jumpy. My mind works all right, but sometimes the words are hard to find to say what I mean, and it sounds as if my mind weren't right. I am more mentally alert in some ways, but in other ways things are harder to grasp. . . . I am easily worried, disturbed and excited. I get afraid sometimes of things that my common sense tells me aren't likely to happen, like my sister getting sick and dying, or something happening to my little girl. If I make an effort I can keep my nerves calm, but it tires me out. I'm terribly, terribly tired all the time—so terribly tired. . . . I haven't wanted to laugh very much, but I can sometimes. But when I try to cry, I can't. It would be a relief if I could. Sometimes my face tightens and I feel as if I had on a mask. I haven't any personality any more. I have changed very much. . . . I have to take Lugol's. I don't

know why I need it but I do. I don't get very depressed, except when I don't take Lugol's. My basal metabolism is never the same. It goes up and down, and doesn't seem to mean a thing. . . . My memory never was good till this disease started. Now it is much better. . . . I never used to be sensitive, but I am now. And the things people say—[sentence unfinished]. . . . I am almost psychic now. I can read what people are thinking and I understand people better. . . . I never used to have much will power but it is much better now. I just can't give up. . . . This disease is worse than having a baby, and a thyroid operation is nothing compared to it."

CONCLUSION

It should be borne in mind that in hypoparathyroidism two factors, at least, are involved: the calcium deficiency and the hormone deficiency. The results of a parathyroid deficiency may be due to any one or more. Those symptoms which are entirely relieved, temporarily, by intravenous calcium gluconate are:

(1) Tetany and consequent physical slowness, poor visual focus, and inability to express emotion.[7]

(2) Tingling sensation.

(3) Numbness, and dullness of the sense of touch.[8]

(4) Increase in the sense of pain.[9]

(5) Extreme cold.

(6) Poor orientation in space and consequent poor body balance.

(7) Mental slowness and confusion, depression and fear.

(8) Loss of initiative.

(9) Slowness in making adjustments and in focusing the attention.

(10) Vibrations.[10]

[7] Hysteria and emotional instability are present only when the blood calcium is unstable, and for a short time afterward.

[8] There is, in addition to the dullness of the sense of touch, a delay in receiving peripheral stimuli. They appear to be two distinct factors.

[9] This is evidently not merely a peripheral occurrence, since intravenous calcium lessens not only peripheral pain but internal pain as well.

[10] The writer believes that the "noises in the head," "roaring in the ears," "vibrations" and "pulsations" are all the same thing; that whether or not the sensation is a sound depends on the rapidity of the vibration. The vibration seems to require an emotional, not a physical, stimulus.

Those factors which are not affected by calcium are:

(1) The extreme sensitivity which causes, in turn, the timidity, the increased observational powers, the increased understanding of others, and the better memory.
(2) Overconscientiousness, and consequent anxiety.
(3) Greater persistence.

These factors are all interwoven. A slow reaction time, for example, is caused by the slowness in the reception of peripheral stimuli and by the lack of resiliency of the muscles, a lack which prevents a quick physical response when the stimulus is received. The poor body balance and stumbling gait seem to arise from the poor vision, the diorientation in space and the dullness of touch. The timidity arises from the extreme sensitivity and to some degree from fear and from the mental confusion.

The writer has no explanation for the difficulty in putting emotions into words. In this respect the disease closely resembles a neurosis, and intravenous calcium serves the purpose of forcing subconscious thoughts and emotions into the conscious mind.

COGNITIVE ASPECTS OF DETERIORATION IN MENTAL PATIENTS *

BY RICHARD S. SOLOMON

The Personnel Institute, Chicago

INTRODUCTION

If we are to have a science of abnormal psychology, there must be a body of experimentally determined facts which are capable of being expressed in quantitative terms. Carney Landis has emphasized such a need in a recent paper in this JOURNAL (9), and, in so doing, has called attention to the work of various investigators who have already made a beginning toward this end. Throughout these investigations, there is the attempt to establish inductively determined categories for the purpose of describing and clarifying the behavior and mental processes of diseased individuals.

In the present investigation, by means of an experimental methodology consistent with the above-mentioned objectives, the writer has sought to ascertain certain basic facts concerning the nature and amount of mental deterioration. In order to achieve this end, it is necessary to think in different terms from those usually employed in the study of mental disorder.

The traditional psychiatric classifications of Kraepelin and others do not lend themselves to experimental investigation. Such conceptual categories as "general paresis," "arteriosclerotic psychosis," or "dementia praecox," are essentially abstractions based upon clinical diagnosis. While they may be valuable terms for purposes of administration and therapy, as research categories they contain too wide an assortment of variables which have no specific reference to any object of an explicitly denotable character. If we think in terms of certain specific and explicitly denotable signs, however, such as the presence of tortuous vessels in the retina, or a four-plus spinal Wassermann, individuals who exhibit such signs (as they naturally occur) can inductively be

* This article was written while the author was a member of the Department of Psychology of Syracuse University; it is based on data collected at the Syracuse Psychopathic Hospital.

placed in mutually exclusive groups for study. The methodology of the following report provides for such a grouping of individuals.

Likewise, the variables of a cognitive nature may be stated in explicitly denotable terms: in terms of the performance of individuals in specific problem situations. There need be no reference to abstract functional unities such as "recent" or "remote" memory. If we can find a number and variety of problems to which individuals may make differential responses, then these responses in themselves *may* serve as distinguishing variables between an individual who, for example, has tortuous vessels in the retina and one who has not. It is thus possible to study deterioration through the application of a variety of specific problem situations.

Each problem situation in this investigation was selected on the basis of diversity. Since we were looking for situations which would most likely give differential responses (in terms of an individual's ability to make a correct response), it was important to choose situations which supposedly involve widely different kinds of operations. It was thought that practicable situations would be those involving the use of (1) general information, (2) word similarities, (3) the reversal of the hands of the clock, (4) the free production of words in relation to articles of clothing, (5) number completion, and (6) disarranged sentences.[1] The term "nature" can be used to refer to the operational differences among the six problem situations.

It will be convenient to speak of the performance of the individuals comprising the various groups in this study in terms of their *capacity level*. Each group of individuals will make a certain score in each of the problem situations; this score may be thought of as expressing their capacity in any of the problem situations at the time of the administration of the problems. Hence, any differences in score in each of the situations between two groups will be considered as indicating a *difference in*

[1] To be sure, the nature of these problem situations was not determined solely on the basis of theoretical operational differences. In a recent investigation Hildreth (8) points out that these are the items "in which the deterioration is not only the most profound, but also persists, as near as can be determined, through every type of pathology and intellectual level." In addition, Thurstone (12) refers to seven elements of intelligence which may be thought of as primary cognitive abilities. These abilities, although they reflect the older interpretation of faculties, bear a striking similarity in their descriptive names to the six problem situations which we have selected.

capacity level between these two groups in that particular situation.[2]

A number of specific questions can be asked concerning deterioration in terms of the performance, in the problem situations, of individuals composing the various pathological groups. Is the *capacity level* the same throughout all situations for the various pathological groups? If not, what significant differences occur? What are the differences or similarities in *capacity level* between psychotic individuals *with* organic pathology, and psychotic individuals *without* organic pathology? Are there any relationships in respect to differences in *capacity level* throughout the pathological groups in the various problem situations? These questions constitute the main problem of this study; other related questions will arise as soon as the problem is analyzed further.

The advantages of the methodology and statement of the problem, given above, will become apparent when the specific findings of the investigation are stated; for then these results can be expressed as relationships or laws, either of co-existence or co-variation, between two variables each of which has an explicitly denotable referent.

THE MATERIALS AND SUBJECTS OF THE INVESTIGATION

The problem situations. Since the items which comprise the problem situations have to be selected primarily for psychotic individuals, it is advisable to present them orally and to score them as rapidly as the responses are given. The number of items cannot be so numerous as to tax the subject's patience, or to fatigue him; but the number of items must be sufficiently numerous so that neither chance nor an unusual familiarity with some particular item will materially affect the score. Then, too, the items should vary in difficulty in order that small degrees of mental efficiency (or deficiency) on all levels, when they occur, will be revealed.

These criteria were fulfilled in a preliminary investigative

[2] There have been a few experimental studies which have been concerned with the problem of intellectual deterioration; but few of these have any direct bearing on the present investigation. Babcock (2) has made a study of deterioration in which the pathological groups were clinically defined so that the findings are in terms of diagnosis. Simmins (10) and Harbinson (6) have both investigated the deterioration of the "G" factor in psychopathic individuals who were symptomatically classified. Their principal interest was in the concept "G" which according to the Spearman technique is determined statistically, not psychologically.

period; a measuring instrument of low variability (as confirmed by future measurements) was then obtained, in each situation. The following are examples of the items comprising each situation; the first is easy; the second, medium; the third, difficult:

I. General Information (20 items):
1. Of what are shoes made?
2. To what country does Siberia belong?
3. What is the function of respiration?

II. Similarities (20 items):
1. knife—fork—spoon
2. book—teacher—newspaper
3. dollar—volt—inch

III. Clock Reversals (10 items):
1. one o'clock
2. quarter after twelve
3. twelve minutes to eight

IV. Production of Words (time limit of one minute).

V. Number Completion (20 items):
1. 8 7 6 5 4 3 —— ——
2. 8 1 6 1 4 1 —— ——
3. 3 4 6 9 13 18 —— ——

VI. Disarranged Sentences (20 items):
1. honey make bees
2. fuel wood are coal and for used
3. clouds rain sky from comes the in

Each item was scored as being passed or failed; a record was kept of all items passed so that they could be analyzed individually or by total score.

The subjects. The subjects were taken in serial order from the admissions to the Syracuse Psychopathic Hospital. Only those patients were included who had been born and educated in the United States, and who were eighteen years of age or older. This restriction was advisable so as to control, as well as possible, any differences in performance due to nativity, education, or maturation.

We have, then, one hundred and five individuals available for study; and for purposes of analysis it is desirable to separate these into various groups. This division was accomplished on the basis of organic pathology by inductive methods. That is, utilizing all of the cases, individuals may be divided into groups as they ordinarily and naturally occur without regard for isolated organic pathology. Separated in this manner, the subjects fall into three groups, as follows:

F group N=30	Individuals without any evidence of organic pathology, but with functional psychopathy.
O group N=41	Individuals who exhibit definite signs of organic pathology such as retinal or peripheral sclerosis or both, positive blood and spinal tests or both, alcoholism, or any combination of these signs.
M group N=34	Individuals who exhibit definite organic pathology, in addition to psychopathy not due to organic origin.

The control. In order to determine that failure of a certain amount on the items in the six problem situations is associated with pathological conditions, we must see how well persons of different degrees of intelligence (within the normal range) can do in these same situations. In other words, what is the performance of individuals who are non-psychotic, non-pathological,

TABLE 1

MEAN SCORES AND STANDARD ERRORS OF THE MEAN FOR THE *N* GROUP AND EACH OF THE THREE PATHOLOGICAL GROUPS FOR THE SIX PROBLEM SITUATIONS

SITUATION	GROUPS							
	N N=50		*F* N=30		*M* N=34		*O* N=41	
	MEAN	σ_m	MEAN	σ_m	MEAN	σ_m	MEAN	σ_m
1. Information	17.04	.29	13.70	.67	13.47	.61	12.24	.73
2. Similarities	16.20	.29	12.63	.93	11.14	1.01	9.75	1.03
3. Clock	7.73	.26	5.20	.53	3.23	.51	4.26	.51
4. Words	19.60	.56	14.90	.76	14.82	.75	14.56	.64
5. Numbers	12.44	.43	10.06	.78	7.00	.68	7.12	.70
6. Sentences	15.88	.34	14.13	.71	11.26	.92	10.12	.92

who could be classified as dull through superior (with the great majority falling in the middle range between 90 I.Q. and 110 I.Q.), and who are comparable to the pathological groups in age, education, social and economic status? Such a group of individuals was obtained; it consisted of 50 cases, and will be referred to as the N group.

DATA, ANALYSIS, AND RESULTS

In Table 1 there are given the mean scores and the standard errors of the mean for the pathological groups, as well as for

the N group, for the six problem situations. These mean scores are total scores; each mean represents the arithmetical average of the cumulative number of items passed in a situation by the individuals comprising a particular group. In respect to these mean scores, we call attention to the fact that in every case the scores of the N group are *higher* than those of the three pathological groups, a fact which, of course, will make all deviations or differences of the means between the N group and each of the three pathological groups favor the N group. Shortly, we shall consider the reliability or the significance of these differences.

Before doing so, however, let us consider the reliability of the total mean scores themselves. It is particularly important to evaluate these scores, for any results obtained from the mean scores can be only as reliable as the means are themselves. Moreover, although variability of the mean score is taken into account when the significance of differences is determined, it is most desirable to know, to a certain extent, that when differences of little significance occur, they are due to comparable performance of the individuals of the various groups, and not to undue variability of the group.

Looking at the standard errors of the mean scores given in Table 1, we see that in every situation throughout every group the chances are equal that any obtained mean score will not deviate from a hypothetical true mean score by more than one item. Such being the case, we are justified in assuming that an adequate and reliable sample of the individuals comprising the various groups has been taken. For if we were to take an infinite number of samples from the same population, we would, just as likely as not, find mean scores of the same number of individual integral total items passed as we have already obtained. (One item is the discrete unit of measurement.)[3]

We are now ready to examine the differences in mean scores between the N group and each of the three pathological groups. These are listed in Table 2; in each case the critical ratio of the difference (based on the standard error of the mean) is also given.

[3] Since we are dealing with a discrete series of items, it is well to point out that any scores representing fractional parts of an item are statistical abstractions, which, although necessary, should be recognized as such.

TABLE 2

DIFFERENCES BETWEEN MEANS OF THE *N* GROUP AND THE THREE PATHOLOGICAL GROUPS FOR THE SIX PROBLEM SITUATIONS, WITH THE CRITICAL RATIO OF EACH DIFFERENCE

SITUATION	GROUPS					
	N–F		N–M		N–O	
	DIFF.	C.R.	DIFF.	C.R.	DIFF.	C.R.
1. Information	3.34	4.54	3.57	5.24	4.80	6.08
2. Similarities	3.59	3.86	5.06	4.82	6.45	6.03
3. Clock	2.53	4.30	4.50	7.81	3.47	6.01
4. Words	4.70	4.95	4.78	5.09	5.04	5.93
5. Numbers	2.38	2.66	5.44	6.72	5.32	6.48
6. Sentences	1.57	2.22	4.62	4.69	5.76	5.87

As we have pointed out before, each difference here is in favor of the N group (cf. Table 1). In order to determine whether these differences are real or significant, or whether, on the contrary, they are due to errors of sampling or other chance factors, we have divided the obtained difference by the standard error of the difference, and have selected quotients of 2.00 or over as indicative of significant differences.[4]

Every difference throughout every situation is significant. In fact, of the eighteen values found in Table 2, there are only two which lie between 2.00 and 3.00; all others are well over 3.00. In other words, the capacity level of the individuals comprising the F, M, and O groups is significantly lower than that of the N group in all of the problem situations.

It is at once apparent that the capacity level is not the same, either in every situation or in any one single pathological group. Of the three pathological groups, the F group consistently shows

[4] Quotients which give values of 3.00 or over are considered to indicate virtual certainty that the obtained differences are significant, since they represent at least .99865 of the total area under the normal probability curve. A value of 2.00, however, indicates very strong probability in this respect, since it represents .9772 of the total area, and is only 2.145 per cent less significant than a value of 3.00. We then have a criterion by which we are fairly certain that there are at least 97 chances in 100 that the differences in mean scores which are obtained are not due to chance factors.

the highest capacity level in every situation; the O group shows the lowest capacity level in four of the six situations; and the M group lies between the two, except in the Clock and Numbers situations, where it is the lowest. In order to determine how much and in what respect the pathological groups differ from each other, we must find the significance of the differences between respective mean scores. These results are given in Table 3.

TABLE 3

DIFFERENCES BETWEEN THE RESPECTIVE MEANS OF THE *F*, *M*, AND *O* GROUPS, WITH THE CRITICAL RATIO OF EACH DIFFERENCE

SITUATION	GROUPS					
	F–M		F–O		M–O	
	DIFF.	C.R.	DIFF.	C.R.	DIFF.	C.R.
1. Information	.23	.25	1.46	1.46	1.23	1.28
2. Similarities	1.49	1.09	2.88	2.09	1.39	.97
3. Clock	1.97	2.67	.94	1.27	−1.03	1.41
4. Words	.08	.07	.34	.34	.26	.26
5. Numbers	3.06	2.94	2.94	2.80	−.12	.12
6. Sentences	2.87	2.45	4.01	3.46	1.14	.87

It is evident from this table that not all differences between mean scores are significant. There are, in fact, no reliable differences between the M and O groups; all reliable differences, of which there are six, occur between the F and M, and the F and O groups. Four of these are common to both sets of differences and occur in situations involving Numbers and Sentences; the other two are unique in each set of differences—the Clock in the F–M groups, and the Similarities in the F–O groups.

There are several findings here that are of particular interest. If we keep in mind that all pathological groups show a significant difference in capacity level from the N group in every situation, it is apparent that there is no significance in the differences in capacity level in any situation between individuals who show signs of organic pathology *plus* functional psychopathy not

known to be due to organic factors (M group), and individuals who show *only* signs of organic pathology (O group). In other words, *the capacity level associated with organic plus non-organic pathology is no lower than the capacity level associated with organic pathology alone.* And it will be remembered that the capacity level associated with non-organic pathology *alone* (F group) was sufficiently low in itself to show significant differences from the capacity level of those individuals who comprise the N group.

In the F–O set of values, although the value for the Similarities situation is significantly high, there is no significant difference in this same situation between the other two groups. We may then infer that a low capacity level in the Similarities situation is especially characteristic of the O group. By similar reasoning it can be determined that a low capacity level in the Numbers and Sentences situations is particularly characteristic of the M and O groups, and that it is comparatively high in the F group; that a low capacity level in the Clock situation is especially characteristic of the M group; and that a low capacity level in the Information and Words situations is not characteristic of any one particular group, nor is it, furthermore, associated with either organic or non-organic types of pathology, since all three pathological groups show the same capacity level in these two situations.

This consideration brings up the question as to whether the differences in capacity level from the N group in the six situations is relatively the same for each of the pathological groups; that is, is there any correspondence between the place each situation occupies in one group and the place it occupies in another group? An inspection of Table 2 yields an approximate solution to the problem, but a more precise answer can be obtained by the use of the rank difference correlation formula. By ranking the values in Table 2 for each group and then taking the difference in rank and applying the formula, a coefficient will be obtained which will express, in some manner, the correspondence between each group in respect to the place each situation occupies in its own particular group. A comparable basis for ranking can be found in the critical-ratio values given in Table 2.[5] The corre-

[5] When these values are used, units of measurement are made comparable in that they represent fractional parts of the total area under the normal probability curve in terms of standard deviations from a given point. In the present case, the mean score of the N group in any situation is the point from which deviations in the three pathological

lation coefficients will indicate, then, the relationship—if it exists—between the difference in the whole array of mean scores of one pathological group from the N group, and this difference in another pathological group from the N group.

These coefficients are: +.54 between the M and O groups; +.31 between the M and F groups; and +.03 between the F and O groups. The relation between the M and O groups is highest; there is no relation between the F and O groups; whereas between the M and F groups the relation is less than it is between the M and O groups. This finding points to the suggestion that certain pathology is associated with the order of magnitude of capacity level throughout the various situations; that the order of capacity level through the various situations in the presence of organic pathology is not similar to that in the presence of non-organic pathology; and that when organic and non-organic types of pathology occur together, the order of capacity level is similar to that associated with *organic pathology alone* rather than to that associated with *non-organic pathology alone*.

These findings have been concerned with the capacity level of each particular group. Now there is the possibility that there are relationships which hold throughout all groups. There may be some relationship between the amount of difference in capacity level among the pathological groups in any situation, and the amount of difference in capacity level from the N group that any two pathological groups show in respect to the same situation. In Table 2 are found, for each pathological group, the difference values from the N group for every situation; in Table 3 are found the difference values among the pathological groups for every situation. On the basis of the critical-ratio values, these differences may be ranked from 1 to 18 according to their magnitude, and the rank difference formula can then be applied. (In the case of the values found in Table 2, the average value for any two groups is taken for each situation.)

groups are taken. It will be seen that the logic for the use of the critical-ratio values (or the difference-equation values, as they may be called) in this connection is very similar to that of the use made of the standard score (X/σ) when individual scores representing different capacities are compared. If it were necessary to scale the items in each situation (a possibility which was very carefully considered), we could easily proceed from this step.

This coefficient of correlation is —.51, indicating that, on the whole, the greatest difference values from the N group tend to go with the smallest difference values among pathological groups; and that, with an assurance of 14 per cent above chance, as the difference values from the N group increase, the difference values among pathological groups decrease; i.e., as the amount of difference of a pathological group from the normal increases, the difference among pathological groups becomes smaller.

In like manner it can be ascertained whether this relationship holds good on an *intra-group* basis; instead of taking differences between groups in respect to each situation, differences are now taken between every situation *within* each particular pathological group and throughout every group, using only the data given in Table 2. There are, then, two sets of forty-five values available for ranking: one set representing the amount of difference from the N group for any two situations; the other representing all possible difference values between pathological groups for all situations. When the rank-difference correlation is applied, the resultant coefficient is —.54, or approximately the same as that obtained above when differences in capacity level were obtained between pathological groups alone.

On this basis it can be stated with an accuracy of 16 per cent better than a guess, that as the amount of difference in capacity level of any pathological group from the N group increases in any situation (taking the groups as a whole), the differences between any two situations within all the pathological groups will become smaller. It would appear that apparently any differences between individuals which are associated with differing types of pathology, or with the same pathology, tend to be equalized as the capacity level becomes lower.

A rough check may be obtained by inspecting the group values given in Table 2 individually. Of the pathological groups, the O group, we know, shows the lowest capacity level, the F group the highest capacity level. By a mere glance at the difference values (critical ratio) for each situation, it will be seen that in the O group they deviate little from each other, whereas in the F group this deviation is comparatively great.

PATTERN RELATIONSHIPS IN THE DETERIORATION OF THE DIFFERING PATHOLOGICAL GROUPS

In the collection of the data we have recorded the specific items passed in each situation for every group. It is therefore possible to find the percentage of the individuals in each group who have passed each particular item. These percentages represent the actual performance on each item of the individuals who comprise each group. The question arises, then, whether the performance is the same for the pathological groups as it is for the normal group in respect to the difficulty of the items. We

TABLE 4

COEFFICIENTS INDICATING DEGREE OF SIMILARITY BETWEEN THE PATHOLOGICAL AND *N* GROUPS IN RESPECT TO DIFFICULTY OF ITEMS IN EACH SITUATION

SITUATION	GROUPS		
	F	M	O
1. Information	.89	.86	.91
2. Similarities	.71	.67	.68
3. Clock	.69	.79	.66
4. Numbers	.93	.92	.92
5. Sentences	.90	.94	.93

take as an indication of the difficulty of each item, the percentage of the individuals of the N group who pass that item. After the *order* of difficulty of each item in every group is determined, the correspondence between the N group and the F, M, and O groups can be obtained through the application of the rank-difference correlation formula.

The correspondence or similarity between the N group and the three pathological groups, or among (as in other cases) the three pathological groups, we may call a similarity of *pattern*. A pattern may then be defined, in accordance with the methodology of this study, as the ranking into which items fall for any pathological group when they are listed in order of magnitude on any particular continuum.

Patterns of difficulty. In Table 4 are found the rank-difference coefficients of correlation which express the degree of pattern similarity between the N group and each of the pathological groups in respect to the difficulty of items. It will be seen that, although there is some variation in the patterns of item difficulty, the coefficients are all positive and high. It is apparent that pathology of these groups of individuals does not affect the order of difficulty and that, although order of difficulty is affected in some situations more than in others, on the whole this order is almost the same for the pathological groups as it is for the normal group. It is evident from Table 4, moreover, that there is very close pattern similarity in respect to order of difficulty of items among the F, M, and O groups; the coefficients are consistently and comparatively high or low in every situation for the three pathological groups. Thus, differences in pathology do not affect difficulty to any great degree in any of the situations.

Patterns of change. The word "change" as we use it simply expresses the differences in item percentages, corrected for reliability, between the N group and each of the pathological groups. *A pattern of change* will indicate whether or not the pathological groups vary alike from the normal *within* each situation. In order to determine this fact, the statistical technique of ranking the items must be employed again. Percentage differences from the N group are computed for each item for each of the pathological groups; these differences, corrected for reliability by Yule's formula for the standard error of the difference between two proportions, are then ranked according to the magnitude of the difference equation. The *similarity of the pattern of change* within each situation between the F, M, and O groups is given by the rank-difference coefficient of correlation. Coefficients approaching +1.00 are considered to indicate strong pattern similarity; those near 0.00 indicate no pattern similarity; and those approaching —1.00 show opposite patterns.

In Table 5 are presented the coefficients of correlation indicating the degree of similarity of pattern of change within each situation among the pathological groups. It is evident from this table that there is considerable similarity in these patterns of change among the pathological groups in most of the situations. There is, nevertheless, some variation between groups in certain situations; also, there is more variation between some groups than

between others, and in certain situations more than in others. One of the most striking features of this array of coefficients is the apparent inconsistency in their variation. In fact, no situation consistently shows either a high or low pattern consistently throughout all situations. From these general observations it may be inferred that apparently similarity of pattern of change within any situation is *not primarily* connected with similarity of the items comprising the situation, or with similarity of pathology of the individuals comprising the differing pathological groups.

In an earlier part of this investigation a relationship was reported between the magnitude of difference in capacity level

TABLE 5

COEFFICIENTS REPRESENTING PATTERNS OF CHANGE WITHIN EACH SITUATION INDICATING DEGREE OF SIMILARITY IN RESPECT TO THE *N* GROUP BETWEEN PATHOLOGICAL GROUPS

SITUATION	GROUPS		
	F–M	F–O	O–M
1. Information	.65	.62	.79
2. Similarities	.62	.71	.76
3. Clock	.57	.68	.22
4. Numbers	.64	.66	.67
5. Sentences	.68	.61	.49

between the N group and the three pathological groups, and the amount of difference between pathological groups. On the whole, as the former became greater, the differences between pathological groups became smaller. It was thought that, as the magnitude of difference in capacity level from the normal became greater in the pathological groups, there was more similarity between the pathological groups in the various situations. This conclusion points to the possibility that there may be a relationship between the similarity of pattern of change within a situation between pathological groups, and the amount of difference from the normal which is shown in that situation.

The simplest way to determine this point is to arrange the pat-

terns (the coefficients in Table 5) in order of similarity. From Table 2 the average amount of change of the pathological groups from the N group in each situation can be determined; and on this basis each situation also can be ranked, and the rank-difference formula applied. When this is done, a coefficient of —.60 is obtained. There is no inherent value to this coefficient, since it is quite removed from the data which have been actually recorded. It does, however, indicate some relationship between the two variables; in short, the situation within which there is the *least* similarity of pattern of change is the situation in which there is the *greatest* amount of difference in capacity level from the normal.

By a similar technique which also gave a rank-difference correlation coefficient of .60, it was found that there is a relationship between similarity of pattern of change from the N group within each situation, and similarity in the *degree* of difference in capacity level from the normal between all pathological groups. It can thus be inferred, to the extent indicated by the coefficient (20 per cent above chance), that the *greatest* similarity of pattern of change within a situation is accompanied by the *greatest* similarity of degree of difference in capacity level of the pathological groups from the normal in that situation.

The relation between difficulty and amount of change. Certain patterns of difficulty and of change have already been indicated. It was found that differences in difficulty of the various items and differences in the change of the pathological groups from the normal have very little association with differences in pathology or the nature of the situations *per se*. Nevertheless, we know that individual items comprising each situation vary in difficulty, and also that there is variation in the amount of change from the N group on each item. There may be a relation between these two variables, a consideration which can again be determined by the application of the rank-difference correlation formula. According to this method the original data are ranked; then, if no relation exists between the difficulty of an item and the change from the N group shown by the individuals comprising each pathological group on that item, the ranks of the items will scatter indiscriminately and the coefficient of correlation will be zero.

In Table 6 are found the rank-difference coefficients of corre-

lation for each situation and each pathological group, showing the patterns between difficulty of items and magnitude of change. Of the fifteen coefficients found in this table, thirteen are negative and two are zero. With the exception of the Clock situation in the O group, the patterns vary consistently from high to low throughout all of the situations. On the whole, it can be said that there is a relation between the amount of change from the N group on a particular item and the difficulty of that item, the relation being such that the *greatest change* is shown on items which are the least *difficult*. Or, in other words, as the difficulty

TABLE 6

PATTERNS INDICATING THE RELATION BETWEEN DIFFICULTY OF ITEMS (*N* GROUP) AND AMOUNT OF CHANGE (FROM THE *N* GROUP) ON EACH ITEM FOR THE PATHOLOGICAL GROUPS WITHIN EACH SITUATION

SITUATION	GROUPS		
	F	M	O
1. Information	—.02	—.16	—.03
2. Similarities	—.57	—.32	—.60
3. Clock	.00	.00	—.66
4. Numbers	—.28	—.24	—.27
5. Sentences	—.58	—.66	—.87

of items *increases,* the amount of change *decreases.* It should be pointed out that this generalization applies to some situations more than to others; and in a few situations, not at all

There is, then, variation in these patterns; and there is the possibility that there may be a relation between this and some other variable. Using the same type of analysis as above, we observe that the Sentences situation has a high pattern throughout the three pathological groups; the pattern in the Similarities situation is also high; yet the patterns in the remaining situations are relatively low. From Table 2 it can be determined that the amount of difference in capacity level of the pathological groups from the N group in each situation (taken as a whole) varies *inversely* with the magnitude of the patterns in the five situations (rank-difference correlation coefficient of —.70, 30 per cent above

chance). In other words, the patterns break down when the change in that situation is great—a fact which might indicate that when the change is small in any situation, the difficulty of items probably plays a discriminating part in the amount of change shown on each item. But when differences in capacity level between the N group and the pathological groups in a situation are great, easy as well as hard items indiscriminately show the same amount of change.

Summary of results. Summarizing the specific results found in this investigation we may state that:

1. The capacity level of the three pathological groups is significantly lower than the capacity level of the N group in every situation.
2. The capacity level in each situation is not the same for the three pathological groups.
3. Of the three pathological groups, the F group shows the highest capacity level in every situation.
4. The O group, on the whole, shows the lowest capacity level.
5. On the whole, the difference in capacity level between the M and the O groups is negligible.
6. There is little difference in the capacity level between the M and the O groups in respect to any one of the situations.
7. A relatively low capacity level in the Similarities situation is especially characteristic of the O group.
8. A relatively low capacity level in the Clock situation is especially characteristic of the M group.
9. A comparatively low capacity level in the Numbers and Sentences situations is especially characteristic of both the O and the M groups; a comparatively high capacity level in these situations is characteristic of the F group.
10. A low capacity level in the Information and Words situations is not especially characteristic of any of the pathological groups.
11. In respect to the differences which the pathological groups show from the N group, the M and O groups are the most similar.
12. In respect to differences from the N group, there is no similarity between the F and the O groups.
13. In respect to the differences from the N group, the M and the F groups show more similarity than do the F and the O groups, but less than do the M and the O groups.
14. As the capacity level decreases, the differences in capacity level between any two groups in any problem situation decrease.
15. As the capacity level decreases, the differences in capacity level between problem situations within all groups decrease.
16. The relative order of difficulty of items is approximately the same for the pathological groups as for the normal group.
17. The similarity of pattern of difficulty between the N group and the pathological groups shows variation to the extent indicated by the pattern coefficients, ranging from .66 (25 per cent above chance) to .94 (66 per cent above chance).

18. There is considerable similarity in pattern of change from the N group between the pathological groups.
19. Similarity of pattern of change between the pathological groups shows variation to the extent indicated by the pattern coefficients ranging from .22 (4 per cent above chance) to .79 (40 per cent above chance).
20. There is a relation between similarity of pattern of change in a situation, and the magnitude of the difference in capacity level in that situation between the pathological groups and the N group.
21. To the extent indicated by a coefficient of .60 (20 per cent above chance), the difference in capacity level of a pathological group from the N group in a situation increases, as the similarity of change in that situation decreases.
22. There is a relation between similarity of pattern of change in a situation, and the similarity of degree of difference in capacity level from the N group between pathological groups in that situation.
23. To the extent indicated by a coefficient of .60 (20 per cent above chance), the similarity of pattern of change in a situation increases, as the similarity of degree of difference in capacity level from the N group between pathological groups in that situation increases.
24. There is a relation in most situations between the relative difficulty of items and the amount of change from the N group in the pathological groups on the items.
25. To the extent indicated by pattern coefficients ranging from .00 (0 per cent above chance) to .87 (51 per cent above chance), as the difficulty of items increases, the amount of change from the N group in the pathological groups on the items decreases.
26. There is a relation between variation in the pattern coefficients of change in respect to difficulty of items in a situation, and the difference in capacity level of the pathological groups from the N group in that situation.
27. To the extent indicated by a coefficient of .70 (30 per cent above chance), as the magnitude of the pattern coefficients of change in respect to difficulty of items in a situation increases, the differences in capacity level of the pathological groups from the N group decrease.

THEORETICAL IMPLICATIONS AND SUGGESTIONS

The theoretical implications concerning deterioration revealed through the foregoing analysis are so wide and varied in scope that all we can attempt to accomplish here is to give some passing indication of their nature.

By referring to statements 1, 5, and 6, given above, it may be seen that any detrimental effect of non-organic pathology is *negligible* in the presence of organic pathology; or, conversely, the detrimental effect of organic pathology is *not increased* by the concomitant presence of non-organic pathology. Both types of pathology are followed by defective intellectual performance. In the organic cases, however, the defect is relatively great, and is thought to be due to actual destruction or necrosis of cerebral

and other tissue. In the non-organic cases, the defect, which is relatively mild, is thought to be due to the interference of other mental trends, conflicts or emotional and affective states of sufficient intensity and duration. If this supposition is true, then in order for inhibitory trends to be effective there must, of necessity, be a sufficient amount of potential intellectual ability to be affected.

This account agrees, somewhat, with the fact that the occurrence of functional psychopathy with feeble-mindedness is not prevalent. It might therefore be that defective intellectual ability finally reaches a point, on a purely organic basis, below which interfering mental and emotional trends do not produce sufficient additional defects, or even cease to operate altogether in this respect.

Consistent with this interpretation are the results summarized in statements 14 and 15. Here it may be seen that, as deterioration becomes progressively lower, the performance of one pathological group becomes more like that of the others. In fact, a point may finally be reached below which an individual's performance (regardless of the type of pathology) in diverse situations becomes relatively the same. In other words, there appears to be *a limit below which differences in intellectual functioning level out, or actually are undifferentiated.*

What we have may be the reverse of the sequence of development. At birth or shortly after, a child and many of the lower animals manifest crude, undifferentiated movements and cognitive functions. Through maturation these general activities become more specialized and specific, and may be described by various so-called "growth" curves. It is possible to add theoretically the other side to these curves and project a sequence of "decline," which may resemble, in part, the rapid acceleration of early development with the "plateau" now placed at the end.

That undifferentiated behavior of this type is associated with a large amount of deterioration is reinforced by the findings given in statements 20–23. These results all point to the fact that any unique performance in a particular situation breaks down in individuals who are reaching a gross state of deterioration. The normal difficulty of any series of problems becomes more haphazard, as do the normal integrated patterns of performance. Apparently original differences in difficulty of the items lose

their discriminative value, as it were, so that problems that were once easy now become just as hard to solve as the most difficult ones.

How, then, can we predict the behavior of the mentally deteriorated individual? While prediction is difficult in relation to other types of responses, in this instance it becomes a relatively simple matter. Regardless of the difficulty of the problem, the type of pathology, or the nature of the situation, grossly deteriorated individuals respond in about the same way. In other words, here is evidence of a type of conforming behavior.

This implication suggests another mode of interpretation which is consistent with those already given. We can think of a continuum of capacity level on which normal individuals may be distributed according to the normal curve of compound probability. Pathological individuals, however, having much lower capacity level with more similarity at the lower end of the distribution, might show a progressively more skewed distribution at this end, as the capacity level became lower. We should then have a steep positively accelerated curve at this terminus of the continuum, or what has been called the J-curve of conformity distribution (1; 3). A result such as this would indicate that some common factor of undue potency was present in the situations which was instrumental in producing a conformity of response.

Further investigation under other conditions into the problem of deterioration is necessary to test the validity of the facts of this study and their interpretations. A study of the behavior of the feeble-minded at various levels would have particular significance. In this respect, Werner and Strauss have recently provided other methods and techniques for the study of the problem in their functional analysis in mental deficiency (13). Still another valuable approach, using controlled laboratory conditions with normal subjects, is suggested by Gellhorn's investigations on anoxemia (5).

BIBLIOGRAPHY

1. ALLPORT, F. H. The J-curve hypothesis of conforming behavior. *J. soc. Psychol.*, 1934, **5**, 141–183.
2. BABCOCK, HARRIETT. An experiment in the measurement of mental deterioration. *Arch. Psychol.*, 1930, **18**, No. 117.
3. DICKENS, M., AND SOLOMON, R. The J-Curve hypothesis: certain aspects clarified. *Sociometry*, 1938, **1**, 277–291.

4. FISHER, R. A. *Statistical methods for research workers.* (4th ed.) Edinburgh: Oliver and Boyd.
5. GELLHORN, ERNEST. The integrated action of the organism exemplified by studies on anoxemia. *Sigma Xi Quart.*, 1935, **25**, 156–165.
6. HARBINSON, M. R. An investigation of deterioration of 'general intelligence' or 'G' in psychotic patients. *Brit. J. med. Psychol.*, 1936, **16**, 146–148.
7. HARRIS, A. J., AND SHAKOW, D. The clinical significance of numeral measures of scatter on the Stanford-Binet. *Psychol. Bull.*, 1937, **34**, 134–150.
8. HILDRETH, H. M. *An experimental study of the relation between certain types of organic pathology and intellectual deterioration.* Unpublished Doctoral Thesis, Syracuse University Library, Syracuse, N. Y., 1935.
9. LANDIS, CARNEY. Methods and conclusions in contemporary abnormal psychology. This JOURNAL, 1938, **33**, 323–331.
10. SIMMINS, C. Deterioration of 'G' in psychotic patients. *J. ment. Sci.*, 1933, **79**, 704–734.
11. TERMAN, LEWIS M. *The measurement of intelligence.* Boston: Houghton Mifflin, 1916.
12. THURSTONE, L. L. A new conception of intelligence. *Educ. Rec.*, 1936, **17**.
13. WERNER, HEINZ, AND STRAUSS, ALFRED. Problems and methods of functional analysis in mentally deficient children. This JOURNAL, 1939, **34**, 37–62.
14. YULE, G. U. *An introduction to the theory of statistics.* (6th ed.) London: Charles Griffin and Co., 1922.

RECENT APPLICATIONS OF *THE A-S REACTION STUDY*

BY RICHARD RUGGLES AND GORDON W. ALLPORT
Harvard University

ELEVEN years ago there was described in this Journal a scale for measuring the tendency of an individual to take the active or the passive role in his daily contacts with his fellows(1).[1] The scale, designed especially for use with the college age, has one form for men and another form for women, and has been employed both by teachers of psychology for purposes of demonstration and by psychological consultants. For users of the scale we are offering the present summary of experience as reported in psychological literature and in private communication to the authors of the test. In addition we are reporting an important revision in the score-values and norms for the Form for Women.

REVISION OF THE FORM FOR WOMEN

For four reasons this Form until now has not been as satisfactory as the Form for Men. (a) Its mean score of +14 was out of line with the mean score of zero that holds in the Form for Men. One undesirable consequence of this situation was that items omitted from the Form for Women spuriously lowered the test score, whereas in the Form for Men omissions had zero effect upon the score. (b) The distribution of scores for women in the earlier form was never "normal," there being too many cases for symmetrical distribution in the range +42 to +32, and too few in the range +32 to +12. (c) The empirically derived scores for certain individual items seemed definitely to contradict the meaningfulness of the scale as a whole and probably resulted from erroneous weightings in the original standardization. (d) The split-half reliability (corrected) for the older Form for Women was only +.77, whereas the corresponding figure for the Form for Men is +.85.

[1] *The A-S Reaction Study* is published by Houghton-Mifflin Co., 2 Park Street, Boston, Mass.

In redetermining diagnostic weights the following method was employed. Two hundred cases were used for the analysis (at first two groups of 100 each, later combined when the reliability was found to be high). The replies of the subjects to each question were cast into quintiles set up according to the total scores of the subjects on the pre-revised Form. The distribution of replies given for Question 4a follows:

QUESTION 4A	REPLY	QUINTILES					TOTAL
		A (MOST ASC.)	B	C	D	E (MOST SUB.)	
A salesman takes manifest trouble to show you a quantity of merchandise. You are not entirely suited. Do you find it difficult to say "No"?	Yes	18	25	29	30	33	135
	No	22	15	11	10	7	65
							200

The scoring weight is a quotient expressing the relative frequency with which the ascendant and submissive subjects select a certain reply. From the above distribution it is obvious that *Yes* is the preponderant choice of submissive subjects, and *No* the preponderant choice of ascendant subjects. The formula employed for determining the weighted score-values is

$$\frac{2}{3}\,\frac{(2A+B)-(D+2E)}{\text{avg. of quintiles}} = \text{Item score}$$

The multiplier 2/3 is an arbitrary figure selected to give the test the same general level of score-values as the Form for Men. Applying this formula to the data for Question 4a we obtain for *Yes* a diagnostic score of —.85, and for *No*, +1.75. Rounded off to the nearest whole numbers, these scores become —1 and +2 respectively.

As a consequence of this re-weighing of all items it was found that five situations or parts of situations in the original scale had no diagnostic value whatsoever. These are accordingly omitted from the revised edition of the test. The Revised Manual (1939) presents all the new score-values and new norms for women. Suffice it here to report the new mean for 600 cases as +4.75 (still not as convenient a figure as —0.01 in the Form for Men) and the S. D. as 20.08.

Figure 1 shows graphically the distribution of scores for the Revised Form, along with the distribution obtained from the same cases according to old-form scores-values. The improvement of the revised form is apparent.

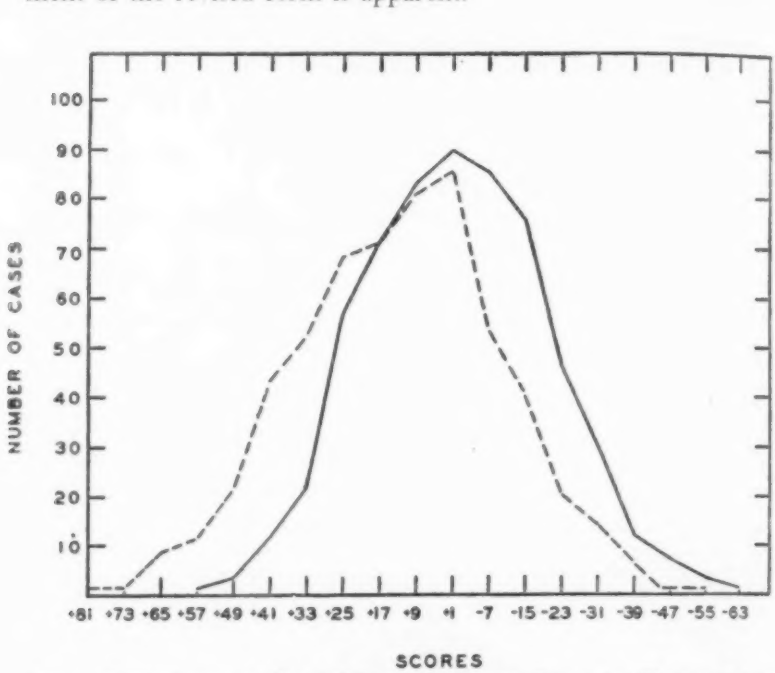

FIG. 1. DISTRIBUTION OF 600 A-S REACTION SCORES, FORM FOR WOMEN, SCORED ACCORDING TO OLD-FORM WEIGHTS (DOTTED LINE) AND REVISED-FORM WEIGHTS (SOLID LINE)

RELIABILITY

Form for Women. For the revised test the uncorrected split-half reliability (100 cases) may be reported as .80. When corrected by the Spearman-Brown prophecy formula this figure becomes $.90 \pm .04$, in contrast to .77 for the previous form. Repeat reliabilities are not available. The correlation between the revised test and the unrevised is $.96 \pm .01$. Though this correlation is high, the improvements in the revised test are numerous enough to justify its replacing the older form.

Form for Men. Various unpublished studies in the authors' hands justify a conservative report on corrected split-half reliability of this Form of .85. This figure might be enhanced

through a pruning of items similar to that reported here for the Form for Women.[2]

Concerning repeat reliability Broom (7) has reported a coefficient of .71 after a three-weeks interval, and Hanna (12) a coefficient of .78 after a one-year interval. Confirming these results several unpublished repeat reliabilities range between .74 and .78. In all cases the intervals were under a year and the subjects were of college age.

VALIDITY

A number of studies, e.g., those of Allport (1), Jersild (15), Broom (7), Symonds (26), seem to show that a suitable figure to represent the correlation between self-ratings and *The A-S Reaction Study,* as well as between associates' ratings and the *Study,* for both the Forms for Men and for Women, is +.45. The range in such studies is considerable, from +.30 to +.70, depending no doubt upon the insight and the training of raters. Since ratings are a notoriously poor arbiter of validity for tests of personality, this figure should not sustain too much weight.

External validation, a more satisfactory method, is indicated in the reports of investigators who have used the scale with selected occupational groups. In all cases the average scores of executives, foremen, and salesmen are distinctively ascendant. The following list summarizes the results of several unpublished studies using the Form for Men.[3]

N	OCCUPATION	MEDIAN SCORE
	REQUIRING ASCENDANCE	
29	Factory foremen	+28
20	General managers of chain stores	+21
21	Branch store managers	+18
29	City executives	+17
64	Factory executives	+14
52	Salesmen in training	+14
83	YMCA secretaries	+12
28	Salesmen	+ 9
	NOT REQUIRING ASCENDANCE	
47	Factory technicians	+ 5
31	Floormen	+ 3
31	Water-meter readers	— 1

[2] Making an item analysis of the Form for Men, Wang (27) reports that 30 items have diagnostic power (i.e., correlate highly with the total score), whereas 11 do not (viz., items numbered 10, 11, 14, 15, 16, 18, 19, 20, 21b, 23a, and 33).

[3] It must be borne in mind that these results are based upon a scale designed primarily for use with college students. Many of the situations are academic in type and therefore not well adapted to use in business and the professions. Because of this limitation R. O. Beckman has revised *The A-S Reaction Study* for use in business (4). His revision can be obtained from the Psychological Corporation, 522 Fifth Ave., New York City.

Schultz (22) reports that a moderate degree of ascendance combined with extroversion and high intelligence is predictive of success in selling. In an unpublished study Maude Williamson finds that though student teachers (female) are unusually submissive, yet a correlation of +.35 obtains between ascendance and their rated success in teaching. Likewise Wang (28) makes an indirect contribution to validation when he reports that he finds associated with ascendance a tendency for participation in games and popularity with one's associates, while a predilection for the study of languages and having limited friendships is associated with submissiveness.

CORRELATIONS

Intelligence. Various published studies (5; 15; 25) and others unpublished agree in finding a zero or negligible relationship between ascendance-submission and intelligence as measured by tests. There seems to be, however, a slight tendency for submissiveness to be associated with high scholarship.

Extroversion-Introversion. That there is a correlation of approximately +.35 between submission and introversion is indicated by several investigators using a variety of extroversion-introversion scales (5; 7; 19; 24; 28).

Neuroticism. That a neurotic tendency goes with submissiveness is the conclusion of investigators who have correlated scores on *The A-S Reaction Study* with scores on the Thurstone *Personality Inventory* (19; 24). That the correlation is spuriously high owing to an overlap of items in the two tests has been shown by Allport (2), who gives the residual correlation between the two scales (all overlap excluded) of +.25.

Persistence. Howells (13) reports a correlation of +.44 between ascendance and persistence; the latter being measured by an experimental situation in which the individual shows his ability to "hang on," even though the situation becomes painful and disagreeable.[4]

[4] The existence of such appreciable correlations as the foregoing raises the question whether ascendance-submission should be considered to be an "ultimate" variable, or whether more basic factors ought to be sought that would account for the recurrent correlations between submission, introversion, neuroticism, lack of persistence, and other "weak" traits. Factor-analysis is now seeking to discover basic underlying variables that will displace these here listed. On the other hand, a case may be made for holding that the correlations here reported are so low and the alphabetic variables produced by factor-analysis so inferior in intelligibility, that for the time being it is wiser to leave the present variables as they stand. (Cf. 3, pp. 242–248, 326–332.)

Dominance. Bernreuter (6) reports correlations of over +.80 between his Dominance scale and *The A-S Reaction Study*. Unfortunately these coefficients are meaningless, inasmuch as his *Personality Inventory* has borrowed such a large proportion of its situations from *The A-S Reaction Study*.

Other variables. Although the *Study* has been correlated with many other variables—learning ability (17), social intelligence (7; 15; 28), suggestibility (28), hypnotizability (30), emotional stability (15), and the like—the results are not particularly illuminating.

GENETIC STUDIES

There have been relatively few studies comparing ascendance-submission at different age levels. Williams and Chamberlin (29) show that there is a steady increase in ascendance during high-school age among girls. Crook and Thomas (10) show that fathers of college students have a mean score slightly higher than these students, while their mothers average slightly lower. The same investigators find no appreciable relation between A-S scores of parents and children, a result that suggests environmental rather than hereditary determination.

The genesis and constancy of ascendance-submission were studied by Sister Mary Aquinas McLaughlin (18). Selecting cases deviating two sigma from the mean in a group of 444 students, she attempted to alter them in such a way as to produce greater "normality." Out of thirteen extremely submissive subjects she was successful in making twelve more ascendant; but only five of her twelve ascendant subjects could she change in the direction of submissiveness. She concluded that ascendance was less modifiable than submission, and that ascendant people are slow in gaining insight, that even with insight they prefer their ascendant role. Sears (23) too has shown that ascendant individuals tend to be lacking in self-criticism.

McLaughlin's study finds that characteristics frequently associated with submissiveness are (a) physical defect, (b) suffering from ridicule, (c) self-comparison with superiors, (d) emotional difficulties, (e) lack of initiative at home. Among the factors related to ascendance are (a) early assuming of responsibility, (b) many social contacts, (c) prowess in athletics, (d) compensation for defects or ill health.

Studying the ascendant and submissive behavior of boys in their homes and at school Schuler (21) shows that although the situational factor is appreciable (certain boys being ascendant in one situation and not in another), yet there is sufficient carry-over to make it seem justifiable to speak of consistency being partly in the *boy* and not merely in the situation.

THE NATURE OF ASCENDANCE-SUBMISSION

Not infrequently *The A-S Reaction Study* is cited as typical evidence for the tendency of personality variables to distribute themselves normally in an unselected population. Figure 1 shows to what extent this is true in the Form for Women. Equally striking is the tendency in the Form for Men (3, p. 333). To have such a distribution is not only convenient, but actually necessary in standardized tests of personality. If such distributions do not occur, the interpretation of norms becomes difficult, correlations with other variables become statistically questionable, and in general the differentiating and diagnostic properties of the test are lessened. Hence testers are pleased to find a continuum that superficially resembles the curve of chance which is duplicated, as we know, in elementary biological measures such as height or weight in unselected populations.

Yet the interpretation of this "normal" curve is by no means simple.[5] In part it is an artifact. In the beginning the authors *sought* a normal distribution of total scores. Figure 1 shows how this manipulation may continue in successive revisions until a more perfect normal curve is achieved. But it is not nature alone that is responsible for the results. Unquestionably, the influence of cultural norms is an additional factor in the distribution obtained. Folkways, civil laws, and other exterior constraints tend to destroy "natural" variation in behavior through their demand for *conformity*. The mode in a distribution is often determined through cultural pressure. It has been shown that conforming behavior usually takes the form of a J-curve.[6] In many situations, a *double* J-curve is found. Some strong social pressure makes for conformity; every individual is expected to fit the same mold, and the majority of them do. But at the same

[5] Cf. J. L. Graham, A paradox in the use of the normal probability curve, *Amer. J. Psychol.*, 1939, **52**, 293–296.

[6] See this issue of the JOURNAL, pp. 419–464.

time individualizing factors enter to cause a certain amount of deviation from the norm, and the tails of the distribution become extended (at a rapidly decreasing rate).

Conceivably as many as four influences may play a part in determining the "normal" distribution obtained for *The A-S Reaction Study:* (a) normally distributed biological variation in some underlying dispositions of temperament (vitality, kinetic impulse, etc.); (b) conformity-producing influences that make for a steeply accentuated mode of behavior neither distinctly ascendant nor distinctly submissive; (c) chance factors in the scale itself that make for a random distribution of scores; (d) the artifact of the authors' intentions, wanting a normal distribution of scores and seeing to it that the test yields it.[7]

However convenient it may be to view ascendance and submission as opposite poles of the same continuum, it must not be forgotten that combining them is an arbitrary procedure. Submission is not merely the absence of ascendance; it is a positive manner of adaptation, as positive as ascendance itself. Many people have a tendency to yield, to take the passive role; it is their *active* style of life. Putting the two complementary modes of adjustment together on a single linear continuum is justified, not because they are one and the same common trait, but because, from the point of view of their adaptive significance, they are complementary. Not one, but two, common traits are represented in the continuum. It is likely that the distribution of either one alone might approximate the J-curve of conforming behavior. Most people do not depart markedly from the norm that is culturally established; social pressure tends to produce moderation.[8]

[7] "The whole problem of the distribution of common traits is exceedingly complex. Though for the present it is well to leave it open, a few suggestions may be made concerning factors influencing the distribution. For one thing the 'normal' curve for common traits does in part reflect nature's preference for average (non-distinctive) levels; but in addition it reflects social pressure toward conformity with some accepted 'average' level of conduct. The normal distribution often results from an arbitrary juxtaposition of opposite modes of adjustment into one linear scale. Finally, the distribution is to a certain extent affected by the inventor of the test who sees to it that successive revisions of his scale yield a more and more symmetrical scatter of scores. There is no serious objection to this last procedure, nor to the normal curve thus obtained, since in any case the scaling of common traits is but a rough way of approaching personality. In the interests of sound theory, however, it is well to realize the extent of the complications introduced into one 'simple' variable for purposes of convenient measurement. Only by so doing can we avoid the fallacy of presuming, as more than one writer has done, that 'everything in personality is normally distributed.'" (3, p. 337.)

[8] Arguments for believing that the "average" scores in *The A-S Reaction Study*

Enough has been said to show that far from clarifying the nature of ascendance and submission in personality, arguments based on the form of distribution of scores actually make for obscurity and superficiality of theory. Ascendance-Submission does not comprise a single, homogeneous, natural, biological variable.

An entirely different approach to the problem may be taken through the study of the evolutionary and genetic significance of ascendance and submission. The investigation by McLaughlin cited above (18) is such a study. According to Jack (14) ascendance and submission can appear as consistent modes of adaptation at an early age, suggesting early formation or temperamental predisposition. Yet striking shifts in childhood also are found when strong enough influences are brought to bear (20). Allport (3, pp. 410–414) discusses the sociological and psychological significance of ascendant and submissive forms of adaptation, and reviews the evidence for considering these forms as relatively self-consistent common traits. It has been suggested by Farram (11) that acendance is identical with the primal urge for power postulated by Adler. In this connection a minor empirical study by Cantril and Allport (8) shows that the relationship is not a simple one. These authors find that individuals with either markedly ascendant or markedly submissive scores are likely to have high political values (i.e., an interest in power as measured by *The Study of Values*). This finding suggests that people who deviate strikingly in ascendance or submission are people in whose lives problems of power and power-relations are especially important and critical.

It is not the purpose of this report to discuss the *power drive,* nor the concepts of *dominance hierarchy, dominance feeling, aggression, leadership, withdrawing reactions,* and the like, all of which in principle are related to the conception of ascendance and submission. Their place in the psychology of personality is an important one, but their ordering and interrelations must wait for the future.

The A-S Reaction Study should, then, be viewed not as an

are the result of true conformity and moderation are given elsewhere (3, p. 336 ftn.). With the data in hand, however, it is not possible to determine whether or not the double J-curve is present in the distribution of scores, because of the influence of other determinants discussed above.

instrument to solve theoretical problems, but as a specific and limited device useful in the coarser analysis pursued by differential psychology. The test itself answers none of the questions concerning the biological or cultural bases of ascendance and submission; it never reveals the highly individual character of ascendant or submissive traits in each separate personality. Like all questionnaire tests it is a forced scale, empirically derived and empirically validated, but only crude and approximate in representing the subtleties of organization in single personalities.

SUMMARY

A revision of *The A-S Reaction Study,* Form for Women, is here reported, together with new data concerning the reliability, validity, and uses of the scale with both men and women. Eleven years of experience have shown that the test is useful for vocational guidance, personal counselling, and psychological demonstrations. The scale defines and measures a bio-social continuum, serviceable for comparative studies within the field of practical psychology, but too coarse in its construction—as are all comparable questionnaire tests—to represent adequately the subtleties of individual traits within the single personality.

REFERENCES

1. ALLPORT, G. W. A test for ascendance-submission. This JOURNAL, 1928, **23**, 118–136.
2. ALLPORT, G. W. The neurotic personality and traits of self-expression. *J. soc. Psychol.*, 1930, **1**, 524–527.
3. ALLPORT, G. W. *Personality: a psychological interpretation.* New York: Henry Holt, 1937.
4. BECKMAN, R. O. Ascendance-submission test—revised. *Person. J.*, 1933, **11**, 387–392.
5. BENDER, I. Ascendance-submission in relation to certain other factors in personality. This JOURNAL, 1928, **23**, 137–143.
6. BERNREUTER, R. G. The theory and construction of the personality inventory. *J. soc. Psychol.*, 1933, **4**, 387–405.
7. BROOM, M. E. Study of a test of ascendance-submission. *J. appl. Psychol.*, 1930, **14**, 405–413.
8. CANTRIL, H. AND ALLPORT, G. W. Recent applications of *The Study of Values.* This JOURNAL, 1933, **28**, 259–273.
9. COOPER, P. Notes on psychological race differences. *Social Forces,* 1930, **8**, 425 f.
10. CROOK, M. N. AND THOMAS, M. Family relationship in ascendance-submission. *Publ. Univ. Cal. at Los Angeles,* 1934, **1**, 189–192.
11. FARRAM, F. Relation of ascendance-submission to neurosis. *Austr. J. Psychol.*, 1935, **13**, 228–232.
12. HANNA, J. V. Reliability of two personality questionnaires. *Psychol. Bull.*, 1931, **28**, 621.
13. HOWELLS, T. H. An experimental study of persistence. This JOURNAL, 1933, **28**, 14–29.
14. JACK, L. M. An experimental study of ascendant behavior in pre-school children. *Univ. Ia. Stud. Child Welf.*, 1934, **9**, 3.
15. JERSILD, A. T. A study of personality. This JOURNAL, 1930, **25**, 115–120.

16. MANZER, C. W. Effect of self-interest on the Allport ascendance-submission scale. *Psychol. Clin.*, 1933, **21**, 268–270.
17. MCGEOCH, J. AND WHITELY, P. L. Correlations between certain measurements of personality traits and of memorizing. *J. educ. Psychol.*, 1933, **24**, 16–20.
18. MCLAUGHLIN, M. A. The genesis and constancy of ascendance and submission as personality traits. *Univ. Ia. Stud. Educ.*, 1931, **6**, 5.
19. MOORE, H. AND STEELE, I. Personality tests. This JOURNAL, 1934, **29**, 45–52.
20. PAGE, M. L. Modification of ascendant behavior in pre-school children. *Univ. Ia. Stud. Child Welf.*, 1936, **12**, 3, pp. 69 f.
21. SCHULER, E. A. A study of the consistency of dominant and submissive behavior in adolescent boys. *J. soc. Psychol.*, 1935, **46**, 403–432.
22. SCHULTZ, R. Tested selected salesmen are successful. *Person. J.*, 1935, **14**, 139–142.
23. SEARS, R. R. Experimental studies of projection. *J. soc. Psychol.*, 1937, **8**, 389–400.
24. STAGNER, R. Intercorrelations of some personality tests. *J. appl. Psychol.*, 1932, **16**, 453–464.
25. STEVENS, S. A. AND WONDERLIC, E. F. The A-S reaction test: a study of Beckman's revision. *Person. J.*, 1934, **13**, 222–224.
26. SYMONDS, P. M. *Diagnosing personality and conduct.* New York: D. Appleton-Century, 1932.
27. WANG, C. K. A. The internal consistency of the Allport's ascendance-submission test (Form for Men). This JOURNAL, 1931, **26**, 154–161.
28. WANG, C. K. A. Personality traits and early history. *Amer. J. Psychol.*, 1932, 44, 768–774.
29. WILLIAMS, G. W. AND CHAMBERLAIN, F. An evaluation of the use of the Allport ascendance-submission test with high school girls. *J. genet. Psychol.*, 1936, 49, 363–375.
30. WHITE, R. Two types of hypnotic trance and their personality correlates. *J. Psychol.*, 1937, **3**, 279–289.

SHORT ARTICLES AND NOTES

EFFECT OF COLLECTIVE LEARNING UPON THE INDIVIDUAL PARTICIPANTS

BY HERBERT GURNEE

Western Reserve University

THERE is evidence that a collective adjustment to a learning situation may be significantly more effective than individual adjustment.[1] By a collective adjustment is meant one in which the trial responses at the choice-points of the situation are determined by group voting. The following questions then arise: To what extent, if any, does this superior performance of the group affect the learning of the individual participants? Is their acquisition facilitated, or is the benefit of participating in the group confined to the result of the collective act as such? Would the individuals have learned as much or more by acting independently? The present article reports an investigation of this problem. Two different kinds of learning situations were used, one employing a maze and the other verbal material.

APPARATUS AND PROCEDURE

Because it is adapted to group situations, a maze of the bolt-head type was used. The panel was a sheet of black fibre, 120 cm. high by 70 cm. wide, fastened to a wooden frame. On the panel were twenty pairs of brass bolts, arranged in two vertical rows and placed 4 cm. apart each way. One of each pair, selected by chance, was wired in circuit with a metal pointer and a six-volt lamp, which was placed on the panel beside the bolt heads; this was the "right" bolt, and the subject's task was to learn to choose it.

The verbal stimuli comprised twenty pairs of two-place numbers, each pair printed one number above the other on a piece of white cardboard. On the other side of the card the numbers were duplicated but in reverse order. By chance one number of each pair was designated "right," and the subject's task was to learn it. The same order of presentation of cards was used throughout, but after each trial one-half were reversed in a prearranged manner to obviate mere positional learning.

There were two groups of subjects of 42 members each, all male undergraduates. The first of these learned alone with only the experimenter present; the others learned in a collective situation consisting of groups of from 9 to 14 members each.

The procedure with respect to the former was as follows. First the maze was presented; then, beginning at the bottom, the subject indicated his choice by an oral response, "right" or "left," for each pair of bolts; and the experimenter did the contacting. Six trials were given, followed by a test trial in which the choices were indicated on a printed diagram instead of orally. With regard to the verbal stimuli, the subject read aloud his choice

[1] Gurnee, H. Maze learning in the collective situation. *J. Psychol.*, 1936, **3**, 437–443.

as each card was presented; if the choice was correct, the experimenter said "right," if not, "wrong." There were again six trials, followed by a test trial in which the responses were written instead of oral.

The subjects in the group or collective situation proceeded in the same way as did those working alone, except that the choice for each pair of bolt-heads or numbers was taken as the majority vote. This vote was decided by the experimenter from the oral responses. A quorum was required to vote, and if the oral preponderance was not clear at any point a revote was taken. After the sixth trial the group members were tested in the same manner as the solitary subjects, writing down their choices instead of giving them vocally.

TABLE 1

DISTRIBUTION OF ERRORS OF INDEPENDENT AND GROUP SUBJECTS ON SEVENTH TRIAL

ERRORS	MAZE		NUMBERS	
	ALONE	GROUP	ALONE	GROUP
0	3	2	2	3
1	0	1	1	2
2	2	2	3	10
3	0	6	6	9
4	5	6	9	9
5	7	7	4	7
6	7	6	4	0
7	4	3	5	2
8	6	5	3	0
9	5	1	4	0
10	1	2	0	0
11	2	1	1	0
12	0	0	0	0
Av.	6.047	5.214	5.024	3.190
S.D.	2.72	2.55	2.56	1.65

RESULTS AND DISCUSSION

The criterion of learning employed was the number of errors on the seventh or test trial. The distribution, means and standard deviations are given in Table 1, for the individuals who worked alone and for those who worked collectively. It is evident that the performance of the latter was superior, both for the maze and the verbal material; which means that, under the conditions of this experiment, individuals will be likely to learn either material better if they operate collectively than if they operate alone. In other words, a social increment may be expected in situations such as these. This increment, obviously greater for the numerical stimuli than for the maze, forces the conclusion that the facilitation, whatever may be its explanation, is in part a function of the kind of situation to which the subjects are required to adjust. The critical ratio for the difference on the numerical stimuli (difference over the standard

error of the difference) is 3.90, which indicates a high degree of significance. With regard to the maze the critical ratio is only 1.46, which falls short of Fisher's test of significance, although it indicates a probability of .928 that the difference is a true one.

There is some evidence that the inferior end of the distribution may have benefited more by the collective situation than the superior end. Thus, on the maze, the same number of individual and group subjects get two or less errors, whereas eight individual in comparison with only four group subjects get more than eight errors. And in the numerical situations, three individual in comparison with five group subjects have less than two errors; whereas eight individual compared to no group subjects get more than seven errors. Interpretation is complicated by the fact that the criterion does not measure overlearning, and overlearning naturally increases as zero errors are approached.

From these data arise two general questions: why is learning better in the collective situation, and why is the collective effect greater for numerical than for maze material? The first query has several possible answers. It may be that subjects try harder in the group, that the presence of other learners arouses competitive tendencies which are absent in the solitary situation. There appears to be no objective method of determining whether this explanation is so. Our only recourse is to verbal reports, and these point definitely in the opposite direction.

There is, however, evidence that working in the group is more pleasant than working alone. In the collective experiment subjects participated almost gaily, in contrast with the serious, tense demeanor of those in the solitary experiment. If pleasure is an aid to learning, as is generally conceded, then it must follow that acquisition would be better in the collective situation, assuming other factors to be equal.

But of probably greater importance were the suggestions to correct response which the group situation afforded. When acting collectively, a subject has, besides the ordinary cues to rightness or wrongness, the observable response tendencies of his fellows, and these may sometimes act to tip his judgment one way or the other, especially in moments of indecision. Since the group was more often right than the average member, this influence may be expected to have worked towards the commission of fewer errors by the participants than would have been the case if they had worked alone; and the result would naturally be a better performance on the test trial. Another factor of importance is the tendency of those who are correct in their choices to react more quickly than those who are wrong; so that doubtful members, being, it is reasonable to suppose, most affected by the earliest suggestions, would be more often led to a right than a wrong response.[2]

To this influence of suggestion on the selection of responses by the individual in the collective situation may be added a factor of social reinforcement or inhibition with regard to the process of impression or learning itself. Thus when an individual responds in a group, the fixation of his reaction will tend to be strengthened or weakened according as that reaction does or does not conform to what he perceives to be the dominant reaction of his associates. That subjects are sensitive to the choices of their associates in such cases is clear both from observations of their behavior and from verbal reports. This consideration ought, then, to mean a more rapid learning rate in the collective

[2] Gurnee, H. A comparison of collective and individual judgments of fact. *J. exper. Psychol.*, 1937, **21**, 106–112.

situation than alone—provided, of course, that effort and other subjective conditions are the same, and provided that the subject is not distracted by the presence of others. There is reason to believe that only a minority are so distracted; and probably as much can be said for increase of effort, although we do not actually know.

The greater social increment for the verbal as over against the maze material is very likely consequent upon the different kinds of reactions involved. It seems reasonable to infer that the reinforcement of the former would be more direct and inevitable, because particularly of the following differences. The maze could be learned in various ways: by orientation with regard to a visual frame of reference, by integration of a pattern of implicit manual reactions or by integration of a verbal sequence. The first seems to have been the predominant method, whereas the overt group responses were of necessity in terms of the last. On the other hand, response to the numbers had to be in verbal terms and so was identical in form with the overt response of the group. Thus whatever effect the group voting had upon the individual member's reactions would be direct and inevitable in the latter case, but not in the former.

SUMMARY

Eighty-four subjects partially learned a bolt-head maze and a series of numerical choices, half of them alone and the other half in a group situation. By test of errors on the seventh trial, those who learned in the group proved to be superior in both the maze and the numbers. The explanation seems to be that those in the group profited by suggestions from their fellows, and that the fixation of their correct responses and elimination of their errors were reinforced by a social factor.

THE DIAGNOSTIC VALUE OF THE VERBAL SUMMATOR *

BY MARY ADAH TRUSSELL

University of Minnesota

EXISTING tests of free association have serious limitations. Word lists as developed by Kent-Rosanoff, while easy to administer and score, are too restricting to allow a wide and therefore significant range of response. The stimulus presented is itself a word, with standard associations, and thus it tends to limit most responses to these associations. Another example is the widely publicized Rorschach ink-blot test. While this test undoubtedly gets closer to the desired "free" association by means of such response categories as movement, detail, color, originality, etc., its limitations have been found to be numerous (**4**). These shortcomings, along with the major one of requiring a highly trained examiner and interpreter, lead us to question the present widespread use of this test for diagnostic purposes.

Skinner (**6**) has suggested a new type of free association test, the verbal summator. In presenting as stimuli neither definite words nor pictures, but vocalized speech samples of a type not comparable to any familiar behavior of the subject, the range of responses is greatly widened and to a greater extent uncontrolled. Repeated presentations of the speech sample bring to the verbalized threshold responses which, in a wide variety of samplings, include the stronger associations of the subject. It should follow from this theory that summated speech is not concerned with the usual external stimuli. Thus the author argues that summation of latent speech, undetermined by other acting stimuli, calls forth vocalized responses which are due to their own relative strength. The value of the verbal summator as a *diagnostic* test must rest, therefore, upon its ability to discriminate between groups of normal and abnormal constitution.

The present exploratory study was undertaken to determine whether or not this technique would yield the suggested diagnostic differences. Indications of its value must be determined before any widespread use can be made of such a method on the basis of claims alone. In order to determine this value, a group of normal men and women was given the test, and the resulting data were compared with those obtained from a group of psychopathic patients. Results from such an analysis should reveal the suitability or unsuitability of the technique as a diagnostic test of free association.

PROCEDURE

The method and apparatus were essentially the same as those used by Skinner, whose article contains a detailed description of procedure, as well as of the summator itself. A modification in this experiment consisted of a

* The author is indebted to Dr. Charles Bird, under whose direction this study was done, and to the helpful suggestions of Dr. B. F. Skinner. Grateful acknowledgment is also made to Dr. J. C. McKinley for permission to use patients in the Psychopathic Ward of the University Hospital.

A more complete account of this research is on file in the Library of the University of Minnesota.

time-saving device for self-recording of answers by the normal group of subjects, helpful also in preventing the subjects from being influenced by previous answers.

The summator is an ordinary electrical phonograph, equipped with a mechanical device, which repeats automatically a given sample until a response is obtained, whereupon the experimenter turns a projecting lever to reach the next sample. Two records of forty samples each were used, one recording having been made by a man, one by a woman, to keep constant any effect of the sex of the speaker. Typical samples are: ēē * ā * and * ō * ī *.[1] Testing was done in sound-proof rooms, with an intensity control to aid in carrying out the illusion of actual words, as revealed to the subjects in the directions. With practically no exceptions, all subjects were easily convinced that the clarity of the recorded speech, rather than any behavior of their own, was being tested. Extraneous hunger stimuli were prevented from appearing by controlled hours of testing. Ages and occupations or educational majors were obtained from official records. The time required to administer the test varied from one half-hour to an hour and a half, depending upon the number of sample repetitions required to elicit a response.

The control group for this experiment consisted of 32 psychology laboratory students, 16 males and 16 females, ranging in age from 18 to 24 years. Of these, 26 were given a retest within 30 days of the first testing. Thirty-two patients in the Psychopathic Ward of the University Hospital in Minneapolis formed the experimental group. Because of varied individual attitudes toward the test, the number of subjects in this group varied with the kind of analysis being made. Age range here extended from 18 to 58 years. It is possible that some of the differences reported below are due to age. Probably, however, most differences between normals and abnormals due to this factor are relatively small, in view of variations within groups and of the nature of the behavior being studied.

In all, the diagnostic classification of the psychopathic patients, assigned in staff clinic only after unusual care, is as follows: 10 psychoneurotics, 5 schizophrenics, 4 manic-depressives—depressed stage, 1 manic stage, 3 paranoids, 2 paretics, 2 psychopathic personalities, 1 toxic psychosis, 1 post-traumatic psychosis, and 2 feeble-minded, one of whom was diagnosed as depressed.

RESULTS

The data permit many comparisons in purely quantitative form, such as the number of samples yielding no response. There are also certain data, such as the content of individual records, which must be treated qualitatively. One of the advantages of this test as a diagnostic instrument, however, is the fact that this qualitative data, once summarized, may be treated statistically and used for purposes of comparison. An example may be found in the treatment of characteristics of the specific psychoses.

Quantitative aspects of the data. A brief resumé will be given of those results revealing little or no statistically significant differences between normal and abnormal groups. A study of the reaction time, as measured by the number of repetitions of the stimulus required to elicit a response, should, according to the theory of Jung and others, reveal the touching of an emotional complex. Means of the number of repetitions required for normal male and female

[1] A neutral sound, indicated here by *, and pronounced u as in *up*, was interposed to differentiate the vowels.

subjects were 12.20 and 11.45 respectively; for male and female psychopathic subjects, 12.63 and 11.42 respectively, *indicating no differentiation of groups.* Retest means for the normals revealed slight practice effects. Measures of variability show for the reaction time a Fisher's F value of 17.79: that is, *a greater variability on the part of the psychopaths.* This was similar to a general trend found throughout the experiment.

A measure was made of the overlapping of responses in the records of the normals, or the subject's two responses to the same stimulus on test and retest. Here a substantial degree of similarity is to be expected, if the verbal summator is tapping latent reserves of strong associations, rather than shifting influences. This measure of percentage of similarity revealed, however, an average of only 5.35: that is, the equivalent of only one word in every four responses of five-word length was found upon retest to be identical.

A consideration of the amount of perseveration shown, undertaken because of its importance in certain psychoses such as schizophrenia, revealed a similar lack of diagnostic discrimination. The definition of perseveration as used by Percival Symonds (7) was employed rather than the more general one used in studies of word frequency; that is, the word appears, perseverates in the next few successive responses, then disappears. Included in the count were perseveration of the word, and of the rhymed words, or sound. The mean number of perseverative responses for each subject was computed, and resulting means for the all-psychotic group were summarized in 5.57 perseverative responses, with a standard deviation of 6.22; the all-normal group yielded an average of 4.38, with a standard deviation of 4.70. *These summator results are in line with those of Hubbard, who found in a study of free-association that measures of perseveration were of little diagnostic use.*

A measure successful in differentiating the two populations was found in the number of meaningless responses given. Jung, and subsequently Kent and Rosanoff, considered such responses as disguises for significant emotional complexes, the real answer being hidden in nonsensical form. In tabulation here, meaningless words were combined with those responses giving only letters, such as "R•M•N•O" etc. The analysis, made in terms of percentages because of psychopathic variability in number of samples responded to, revealed a mean characteristic percentage of meaningless responses of 8.59 for the 32 normals, with a standard deviation of 13.43. The mean percentage for the 29 psychopaths, weighted for difference in number answered, was 18.62, over twice as many, with a standard deviation of 26.34. A Fisher's Z applied to the difference between standard deviations yielded an F value of 3.819; therefore, *the psychopaths tend to have significantly more meaningless responses, and vary among themselves more than the normals.*

Analysis of content by groups. Among the many data analyses made in regard to the thought content of responses, four were felt to be of primary importance. Of these, two measures revealed no discriminatory power whatsoever. The first, the number of references to the subject's occupation, yielded the small totals of 12 and 9 references respectively for normal and psychotic males, 4 and 3 respectively for normal and psychotic females. *Thus no differentiation between preoccupation with emotional ideas, as opposed to everyday life, was found.* Rather, there is the fairly natural difference favoring the entire male sex.

A more complicated survey was next undertaken, that of the manifested egocentricity of the responses, which might reveal the greater personalization

of abnormal delusions, hypochondria, and wishful thinking. Content was classified for each subject under responses involving *I, we, us, me, mine,* or *our,* percentages again being used because of differences in the number of samples completed. The mean percentage of egocentric answers for the 32 normals was 19.30, with a standard deviation of 22.2, and for the 29 psychotics, 18.60 with a standard deviation of 24.0, *revealing no discrimination by means of the amount of egocentricity found.*

According to the hypothesis behind the verbal summator, an individual would be expected to reveal traces of an emotionally-toned complex or idea by frequent references to it in his responses. Summated speech should be concerned with the stronger associations of the subject, therefore, rather than with the usual external stimuli. In testing this assumption, a computation of the total number of references to the test itself, to the experimenter, and to the room, revealed the small totals of 29 by the normals, 23 by the psychotics. Thus we find that *immediate environmental factors play a small part in the response to the samples.*

A more important indication of significant latent content of responses is to be found in their thematic content. Accordingly, all of the records were analyzed in terms of the number of references made by each subject to a specific theme, on the assumption that the stronger complexes would be made evident by more references to them. Responses were then divided into two groups: the total number of subjects referring to a theme over four times, and the total number of these themes. Resulting data show 46.8 per cent of the 32 normal subjects referring to 5 or more themes, an arbitrarily chosen standard, as against 83 per cent of the 26 psychotics. The total number of themes referred to by these normals was 15; by the smaller group of psychotics, 27. It is thus apparent that *psychotics are concerned with relatively fewer revealed ideas, but these specific ideas occupy their attention more completely and intensely than is the case with the normals.*

Analysis of specific psychotic characteristics. Differences between the psychoses form one of the most important considerations in an evaluation of the verbal summator as a diagnostic tool. The most direct way of testing this function seemed to be that of comparing the responses of a specific psychotic group not only with the characteristics commonly attributed to that psychosis, but also with the characteristics of the other psychoses. Although ten psychotic groups were sampled, only the three containing the largest number of subjects were chosen, *viz.,* psychoneurosis, schizophrenia, and paranoia. Lists of psychotic characteristics commonly accepted and measurable by verbal behavior were obtained from textbooks and psychiatric authorities. Because of differences in the number of samples completed, analyses were made in terms of total number of responses for each group.

It is evident from a study of Table 1 that the verbal summator discriminates the psychoses here compared in some of the characteristics measured, but not in others. There seems to be, for example, a definite discrimination of the neurotics from the schizophrenes and the paranoids with respect to references to fears and anxiety, and with respect to responses given in the form of a question. The differences found in the amount of paranoid egocentricity and inferiority might also be considered significant. On the other hand, the amount of stereotypy (thematic perseveration) and formal perseveration found in the paranoid's responses are more typical of expected schizophrenic behavior than that revealed by these schizophrenes themselves. In view of these findings,

it seems safe to conclude that, *as measured here, the verbal summator has little value in discriminating between specific psychoses.*

As a final analysis, an examination of those subjects refusing to undertake or to complete the test should be of value. In this experiment all of the normal group were willing to answer the 40 samples, but 8 psychopaths out of 32, or one-fourth of the total number, failed to complete all the samples. Of these, the two who refused to try the test at all had been diagnosed as in the depressed stage of manic-depression, and the other as a schizophrenic. Of the remaining five refusals, all attempted from four to twenty-eight samples, but usually left many with no response. It was seen that, as might be expected, refusals occurred mainly in those psychoses characterized by apathy—for example, three depressed patients and two schizophrenes. It seems significant that none of the ten cases of psychoneurosis was unwilling to be tested.

TABLE 1

COMPARISON OF GROUPS WITH RESPECT TO THE ALLEGED CHARACTERISTICS OF SPECIFIC PSYCHOSES

CHARACTERISTICS	RESPONSES OF THREE CLINICAL GROUPS					
	NEUROTICS NUMBER=351		SCHIZOPHRENES NUMBER=140		PARANOIDS NUMBER=120	
	N	Per cent	N	Per cent	N	Per cent
NEUROTIC:						
1. Anxiety and morbid fears	55	22	13	9	19	15
2. Hypochondria	17	7	1	1	4	4
3. Question form of response	26	10	2	1	3	3
4. Doubt and self-criticism	5	2	3	2	6	5
5. Egocentricity	56	22	32	23	32	27
6. Perseveration (moderate)	38	15	7	6	25	21
SCHIZOPHRENIC:						
1. Perseveration (marked)	38	15	7	6	25	21
2. Bizarre and absurd responses	30	11	23	16	18	15
3. Telescoping	5	2	5	4	1	1
4. Ideas of reference (grandiose)	3	1	2	1	3	3
5. Possibly over-religious	3	1	0	0	0	0
6. Stereotypy (thematic perseveration)	82	35	34	24	53	44
PARANOID:						
1. Persecution	23	9	8	6	8	7
2. Egocentricity	56	22	32	23	32	27
3. Inferiority	14	6	5	4	9	8
4. Wishful thinking	24	10	19	14	5	4

CONCLUSIONS

As was previously stated, this study was undertaken, not as a conclusive test of the diagnostic value of the verbal summator, but as an exploratory survey to determine certain trends indicative of its value. The results found

are, therefore, not to be taken as the final evidence. Certain conclusions as revealed by these data can, however, be made.

1. The records of these abnormal subjects can be differentiated from the normal on the basis of the number of samples to which no response was given, the number of meaningless responses, and the number of syllables in the response. Of little or no diagnostic value according to these data are the amount of perseveration and the mean number of repetitions required to elicit a response.

2. An analysis of the content of responses reveals no difference in references to occupations and no diagnostic value of the mean percentages of "personalized" or egocentric responses of the subjects.

3. There are two results which bear out the hypothesis stated at the beginning of the paper, that is, the value of the summator in reaching latent verbal content and emotionally-toned ideas through the use of unfamiliar stimuli. These are the lack of reference to external stimuli, and the thematic content as measured by the number of references to special themes. In the latter analysis, the abnormal group revealed a greater average number of themes referred to more than four times, and a greater individual frequency of appearance of these references.

4. A noticeable trend revealed by the data was the significantly greater variability shown in general by the psychotics in the characteristics studied.

5. The data failed to reveal any sex differences in those traits measured, in either the normal or the abnormal group.

In this study the data have been interpreted merely in terms of their clinical significance, with no consideration of the value of the verbal summator as an instrument of research in other types of investigation. Although few qualities of summated speech have here been found to differentiate abnormal from normal subjects, this technique doubtless has value in other directions.

BIBLIOGRAPHY

1. BECK, S. J. *Research Monograph*, No. 1, American Orthopsychiatric Association, 1938.
2. BLEULER, E., AND BRILL, A. A. *Textbook of psychiatry*. New York: Macmillan, 1924.
3. HENDERSON, D. K., AND GILLESPIE, R. D. *A textbook of psychiatry*. New York: Oxford University Press, 1933.
4. HERTZ, M. R. The Rorschach ink blot test. *Psychol. Bull.*, 1935, **32**, 33–66.
5. MORGAN, J. J. B. *The psychology of abnormal people*. New York: Longmans, Green, 1928.
6. SKINNER, B. F. The verbal summator and a method for the study of latent speech. *J. Psychol.*, 1936, **2**, 71–107.
7. SYMONDS, P. *Diagnosing personality and conduct*. New York: Century Co., 1932.

A NOTE ON THE RELATION BETWEEN DIVERGENT INTERESTS AND EMOTIONAL STABILITY

BY RUDOLF PINTNER AND GEORGE FORLANO

Teachers College, Columbia University

It has been suggested that emotional instability may be a concomitant of conflicting interests or of extremely intense interests in the individual. In the present investigation we have taken the Allport and Vernon Study of Values Test as a measure of interests and the Thurstone Personality Schedule as a measure of emotional stability. Our population consists of 240 women university students.

Our first approach was to consider the two highest interest values of each individual. This we did in order to raise the question as to whether individuals whose two highest interests were, for example, in the theoretical and social areas, were more or less emotionally stable than the normal group—i.e., the total group. A similar comparison was made for every other paired combination of the six interest values of the Study of Values Test. If, among the two highest interest values of any individual, two values were equal in score, both were used. In this way we have 315 comparisons for 240 individuals.

These various combinations are shown in Table 1. There were, for example, 28 cases where the two highest interest values were theoretical and social; 9 cases of highest interests in theoretical and political; and so on with all

TABLE 1

MEAN PERSONALITY SCORES OF VARIOUS INTEREST GROUPS

HIGHEST INTERESTS	N	MEAN SCORE	STANDARD RATIOS	
Theoretical and Social	28	38.04	.77	.86
Theoretical and Political	9	38.06	.44	.45
Theoretical and Aesthetic	29	37.16	.90	1.02
Theoretical and Economic	8	28.75	2.82	2.91
Theoretical and Religious	11	43.41	—.39	—.41
Economic and Aesthetic	20	38.00	.61	.62
Economic and Social	23	38.81	.54	.59
Economic and Religious	13	35.20	1.18	1.24
Economic and Political	12	42.09	—.19	—.20
Aesthetic and Social	38	43.03	—.57	—.66
Aesthetic and Political	12	44.59	—.46	—.47
Aesthetic and Religious	44	39.99	.31	.43
Social and Religious	44	35.46	1.72	2.06
Social and Political	11	51.14	—1.48	—1.55
Political and Religious	13	47.12	—1.06	—1.12

the other combinations of the six interest values of the test. The number of cases for each of these various combinations varied from 9 to 44. The most frequent combinations were aesthetic-religious and social-religious; least frequently combined with other interests was political interest. The reason was that our group of women students in education showed high aesthetic, social and religious interests but low political interests.

Column 3 of the table shows the mean scores of the various groups on the Thurstone Test. The next step was to subtract these mean scores from the mean of the normal group, i.e., of the total group of women students, which is 41.13. The differences between this mean and each of the means of the various groups shown in Table 1 were then expressed in terms of the standard error of the difference between the means, as shown in column 4. Nearly all of these standard ratios are negligible. Finally, we calculated new means for our total group by omitting each time the individuals in each of the subgroups with which the total group was to be compared. The standard ratios of the differences between these means are given in column 5.

TABLE 2

SAMPLE	N	METHOD I	METHOD II	METHOD III
1	25	+.24	−.16	−.16
2	25	+.14	+.03	+.14
3	25	+.02	+.12	+.13
Total	75	+.14	−.02	−.05

A high score on the Thurstone indicates emotional instability. Any tendency for these high conflicting interest-value groups to show greater than average emotional instability, therefore, should be indicated by minus signs for our standard ratios. In the present instance this does not turn out to be the case: we find both negative and positive signs. None of the ratios with minus signs is statistically significant. The two largest ratios are positive, which means that the two groups in question are more emotionally stable than the whole population on the Thurstone Personality Schedule. There is obviously no tendency for any combination of two high interests to deviate markedly in emotional stability from the group as a whole.

Our second approach was to inquire whether deviation of interest-value scores from the central tendency of the group or whether deviation of the individual's own interest-value scores from the central tendency of his own scores was correlated with emotional stability as measured by the Thurstone Personality Schedule. For this purpose we gave each subject three scores as follows:

Method I. The standard deviations of the six interest values for the total population were obtained. The subject was given a score of one for each of his six interest values which fell above or below one sigma from the mean of the total population.

Method II. The deviation of each of the six interest values from the mean of each such value for the total population was computed for each subject. His score was the median of these six deviations without regard to signs.

Method III. The score here was the median deviation of the six interest values from the median of the subject's own six interest values.

The scores obtained by these three methods were correlated with the scores on the Thurstone Personality Schedule. We tried three random samples of 25 cases, and finally combined all these samples into a total group of 75. Since all our correlations were practically zero, it did not seem profitable to continue any further with larger numbers. The correlations are shown in Table 2.

There seems, therefore, no hint of any suggestion here that individuals whose interest values deviate markedly from the general population to which they belong, or individuals whose interest values deviate markedly from their own central tendency, show any greater or less emotional stability, as measured by the Thurstone Personality Schedule.

There have been many theories proposed to explain the cause of emotional instability. One of these hypotheses regards the condition of neuroticism as a result of a period of conflict in strong fundamental impulses. Certainly the interest values as measured by the Allport and Vernon test can hardly be characterized as strong and fundamental; in so far as these interest-value scores represent the degree of interest in the various value areas, no relation was found between divergent interests and emotional instability. It may very well be, however, that different results would be obtained with measures of the more elemental impulses or appetites.

REVIEWS

EDITED BY EDNA HEIDBREDER

GENERAL PSYCHOLOGY FROM THE PERSONALISTIC STANDPOINT. By William Stern. (Trans. by H. D. Spoerl.) New York: Macmillan, 1938. Pp. xxii+589.

The revolt away from the theological and metaphysical doctrine of the soul led logically enough in one direction towards a pure phenomenology, a psychology in which mental events stand utterly unrelated to any substrate. Such a program, however, was never carried out. The "stream of consciousness" became a quasi-substratum for the mental events flowing in the stream. And positivistic psychologists more and more rested their science upon the substrate of a nervous system—quite mythically conceived.

It is significant that the difficulties with this "nervous system" have not turned us once more towards phenomenology but towards a more generous concept, that of the "organism," as the basis of mental life. We seem unwilling, if not unable, to conceive event without substance, function without structure.

It has been customary, however, to assume, quite naïvely, that the organism can be described in purely biological, even in purely physical, terms—apparently because the term *originated* in biology. Against this custom Stern's treatment, as set forth in his *General psychology from a personalistic standpoint,* is a dignified protest. The nature of organism can be known only from the phenomena it is postulated to explain or support. Stern thus asserts that the person—which is his term for what many Americans have been calling the organism—is psychophysically neutral.

"Neutrality," however, is too negative an expression for his real meaning. The person, to be sure, should not be thought of in terms of the attributes of either physiological or psychological functioning—in that sense it is neutral—but it is a positive and dynamic reality which accounts for the events of human life, both the "vital" and psychological. (Indeed, Stern recognizes a third category, the evaluative activity which is super-psychological.)

That dualism of substance which stems most clearly from Descartes is thus resolved by Stern in terms reminiscent not only of Spinoza (to whom Stern refers) but of Aristotle: one substance, the person, manifests qualitatively different activities of three levels—the vital, the experiential, and the evaluative. Such a conception in itself is not inacceptable to most "tough-minded" psychologists who, as we have said, simply assume that the one substance is exclusively physical. It is so far acceptable also to a large group whom we may call the incurious (among them most parallelists or double aspectists), who are content to define the substratum as at once physical and psychical but hold that further specifications as to the nature of the person are inadmissable.

It is at this point that Stern's metaphysical position takes an empirical turn. Only by a sterilizing abstraction, he holds, can the phenomena of mental activity be described without reference to their place in the total structure of the person. Psychology cannot rest content with the merely formal assump-

tion of a person as substrate but must actively investigate the nature of the person, must recognize the inherently personal nature of experience.

There is thus posed a specific challenge to personalism: show us in detail that facts are thus differently and better described and systematically interpreted. What new categories of description emerge when account is taken of the person and of what value are these categories? Do they bring new facts to light? Suggest new experiments? Enable us to control the course of events more effectively by using these categories?

The whole of Stern's book may be taken as his answer—the culminating work of one of the greatest scholars and productive thinkers in the psychology of our day. A reviewer's judgment of such a large-scale enterprise inevitably reflects his own biases. The present reviewer has for many years advocated a personalistic orientation in psychology. He has contended, moreover, that the next great advance in psychology waits upon our discovery of new descriptive categories. It is, therefore, with deep regret that he finds himself returning a Scotch verdict of "not proven" for Stern's monumental work.

By this judgment we by no means imply that the attempt to describe experience personalistically has failed. Many of Stern's concepts, we believe, will survive; which ones, however, is not clear. Let us consider one example.

Memory, Stern holds, is a conspicuous instance of a concept which is meaningless without *explicit* and detailed reference to the personal substrate. Memory not only depends upon an enduring structure (as is recognized in the doctrine of mneme or of "traces") but its *character is determined* by the nature of the relation of the learning and recall experiences to the person. Thus Stern finds that there is a continuous series between memory traces which are deeply embedded within the person and those which are more superficial and less firmly attached. The latter are often spoken of as salient, as emerging from the totality of the person. Again we have the related distinction between "bound" and "free" memories. In similar fashion we find reactive thinking related to the "salient," while spontaneous thinking is more deeply embedded. Indeed the contrasting terms *salient* and *embedded* are clearly prime favorites which do duty in many contexts.

And that is just the trouble. One should not be disturbed that the terms are patently metaphorical. All descriptive terms start life as metaphors. But the fact that we can pass so readily from one set of contrasting pairs of descriptive terms to another set which is at least partially synonymous seems evidence that the subordinate properties or attributes denoted by these terms have not been clearly enough envisaged. One has the feeling that Stern has a distinction—or more often a whole set of distinctions—which for all his subtlety he cannot quite reduce to concrete terms.

To the reviewer it is clear enough why this is so. Dialectics are necessary in any science, and psychology has suffered no little from its revolt against self-conscious dialectics; but they are no substitute for a wealth of empirical and experimental fact. Stern has been in too great a hurry—even though this is truly a life work—to build a consistent system without waiting for the patient modeling and reshaping of each block by factual investigation. Although Stern claims that his distinctions rest upon empirical fact and experimental investigation, explicit reference to experiment is rare. Even more serious is the fact that Stern seems little concerned to state his problems in such wise as to challenge experimental confirmation. In short, the nice balance between investigation and systematization has in this work never been even approxi-

mated. The result is at least a relative sterility. The chief value of the book is its challenge to the complacency of the orthodox. For however vaguely he has described them, Stern has called attention to a whole series of dimensions of experience which can—apparently—be successfully treated only by a personalistic psychology. And one fears that the complacently orthodox will never have the patience to read his book.

H. B. ENGLISH.

Ohio State University.

THE CLINICAL TREATMENT OF THE PROBLEM CHILD. By C. R. Rogers. New York: Houghton Mifflin, 1939. Pp. 393.

This book represents a distinct departure from ordinary methods and a significant advance in devoting ten of its twelve chapters to the comprehensive and intensive consideration of methods of treatment and their results; that is, it takes up where a number of other useful books in clinical psychology stop. Although only two chapters are devoted specifically to diagnosis, nevertheless the methods and results of diagnosis are implicit throughout all later discussion. Major attention is given to the kinds of treatment, to the considerations determining the nature of treatment, and to the results of treatment.

The author views treatment as comprising all those means, regardless of the professional person responsible, found helpful and effective in changing the behavior or adjustment of children. By diagnosis he means the determination of the dynamic relation between symptoms and causes, rather than mere classification or description of symptoms.

The first two chapters include brief discussions of several methods of diagnosis—personality tests, the ego-libido method, the case history method—their spheres of usefulness and their limitations. In the third chapter the author's component-factor method of diagnosis is presented. The appendix of the book includes instructions for using this method. Briefly stated it consists of the rating (by means of sample ratings) of the forces operating in the child's experiences classified under eight headings: hereditary; physical; mental; family influence; economic and cultural; social; education and training; self-insight. The method, which aids in focusing treatment on those elements which are alterable, permits predictions as to the direction and extent of the influences resulting from treatment. It represents a step in the direction of providing an objective check on diagnostic thinking and therefore a step away from the "intuitive diagnosis." This emphasis is fundamental throughout the book.

After the presentation of his method of diagnosis, the author launches into the consideration of treatment under several major headings: treatment through change of environment, i. e., through foster home and institutional placement; treatment through modification of environment, i. e., through change in parental attitudes and child-parent attitudes, through the cooperation of the school, through participation in social groups and attendance at camp. All the factors which determine the nature of treatment in the particular case are discussed with exemplary thoroughness. The viewpoint with respect to the specificity of the treatment in the particular case may be illustrated by the following excerpt. "The psychiatrist who would be shocked at the thought of prescribing 'some kind of sedative' for a child will advise 'some kind of recreational outlet' and feel no twinge of conscience." These chapters on change and modification of environment are pertinent especially to the so-called under-privileged

child; they deal with the types of problems more often handled by the social agency as opposed to the problems of private practice.

The last section of the book is devoted to treatment dealing with the individual. Treatment interviews and techniques, educative and expressive techniques, the "deeper" therapies—psychoanalysis, relationship therapy, and other therapies—are discussed with a judicious appraisal of their value, scope, and applicability. One of the outstanding merits of this book is the author's ability to judge dispassionately both the positive value and the limitation of various methods and philosophies of treatment. This statement holds true whether the limitations are both obvious and serious, or whether they are minor; whether the author is evaluating his own work or that of others.

Some of the important topics discussed in this section are the following: the qualifications of the therapist; the factors which determine the applicability of a particular therapy; the use of the personal influence of the therapist; the treatment relationship in the various therapies and its termination; the goals of therapy and the importance of insight and re-education in attaining the therapeutic goals.

Throughout the book the author draws richly on his own clinical experience for illustrative cases and material. In addition, he has brought together many important descriptions of and evaluations of methods and results of treatment, a compilation of great value to any clinical worker. Bibliographies are particularly comprehensive.

The book closes with an important chapter on considerations connected with the nature of the clinic set-up, inter-agency responsibility and cooperation in treatment, and the role of the different professional workers. Significant points brought out in connection with the latter are the overlapping of the contributions of different clinic and community workers, the varying roles given to the different professions in different clinic organizations, the fact that no one professional group is adequately trained to do an all-round job of diagnosis and therapy, and the fact that the function of a staff member at the present time is determined to a considerable extent by his individual experience and aptitude and to a less extent by his professional label.

The author points to the eventual birth of a new professional group with training drawn from the fields of medicine, psychology, sociology, social work, and education. The reviewer would underline the author's statement as to the importance of individual experience and aptitude in determining the worker's function, and would suggest that recognition of this fact is of major importance in the growth of such a new group.

The Clinical Treatment of the Problem Child is an indispensable book to every clinical or community worker, to every student of child and clinical psychology, and to his instructors in these subjects.

ROSE G. ANDERSON.

Psychological Service Center
New York City.

NEW WAYS IN PSYCHOANALYSIS. By Karen Horney. New York: W. W. Norton and Co. Pp. 313. $3.00.

Those who have read *The neurotic personality of our time* will recall that in that volume Dr. Horney, a practicing psychoanalyst, offered certain criticisms of a number of the orthodox Freudian interpretations, and presented her own theory of the structure and causation of the neurotic personality. The book

was highly praised for its bold and intelligent attempt to give psychoanalytical theory a sounder sociological orientation. *New ways in psychoanalysis* is essentially a detailed elaboration of the themes introduced in the former work.

Dr. Horney sets herself a mammoth task—a task at which one could hardly hope fully to succeed at this stage of psychological thought. A quotation from the introduction reveals clearly the purpose of the book:

> Thus the purpose of this book is not to show what is wrong with psychoanalysis, but, through eliminating the debatable elements, to enable psychoanalysis to develop to the height of its potentialities. As a result of both theoretical considerations and practical experience, I believe that the range of problems which can be understood is enlarged considerably if we cut loose from certain historically determined theoretical premises and discard the theories arising on that basis (p. 8).

A passage immediately following the one just cited gives the thesis of the book:

> My conviction, expressed in a nutshell, is that psychoanalysis should outgrow the limitations set by its being an instinctivistic and a genetic psychology. As to the latter, Freud tends to regard later peculiarities as almost direct repetitions of infantile drives or reactions; hence he expects later disturbances to vanish if the underlying infantile experiences are elucidated. When we relinquish this one-sided emphasis on genesis, we recognize that the connection between later peculiarities and earlier experiences is more complicated than Freud assumes: there is no such thing as an isolated repetition of isolated experiences; but the entirety of infantile experiences combines to form a certain character structure, and it is this structure from which later difficulties emanate. Thus the analysis of the actual character structure moves into the foreground of attention.
>
> As to the instinctivistic orientation of psychoanalysis: when character trends are no longer explained as the ultimate outcome of instinctual drives, modified only by the environment, the entire emphasis falls on the life conditions molding the character and we have to search anew for the environmental factors responsible for creating neurotic conflicts; thus disturbances in human relationships become the crucial factor in the genesis of neuroses (pp. 8–9).

After thus stating, and elaborating, the purpose and thesis of the book in an introductory chapter, the author sets herself earnestly and systematically to the task of fulfilling that purpose and establishing that thesis. Chapter I is called "Fundamentals of Psychoanalysis." In discussing the difficult and highly debatable problem as to what constitutes the basic principles of Freud's psychology Dr. Horney says: "I regard as the most fundamental and most significant of Freud's findings his doctrines that psychic processes are strictly determined, that actions and feelings may be determined by unconscious motivations and that the motivations driving us are emotional forces" (p. 18). Again: "Not least in importance, Freud has given us basic methodological tools for therapy. The main concepts which have contributed to psychoanalytical therapy are those relating to transference, to resistance and to the method of free association" (p. 33). These concepts are offered among others as the relatively undebatable fundamentals of psychoanalysis. Conspicuous for its absence from the fundamentals is the psychosexual theory of personality development—that is, the whole structure of the libido theory. This chapter thus raises a nice issue for psychoanalysts, namely, what is the essential test of fellowship in the psychoanalytic brotherhood. Fortunately, Dr. Horney does not seem to have the slightest interest in that question.

In Chapter II an attempt is made to delineate the premises on which the Freudian theories have been built, and to show that even a mind of such excellent quality as Freud's could not escape being influenced by the mentality of his time. The argument runs that the setting in which Freud's thinking grew caused him to accept more or less unconsciously certain premises upon

which his thinking rests, and from which it frequently gets a theoretical bias that determines its direction. His most basic interpretations have their roots deep in certain 19th century philosophic biases. For example, Freud's whole system of thought is affected by his "biological orientation" accompanied by a very nearly complete innocence of recent sociological, cultural, and anthropological findings. Out of this unbalanced emphasis comes a corresponding emphasis upon heredity and a neglect of environment as forces which shape the personality structure. Other premises are similarly examined.

Of the fourteen remaining chapters of the book one is given to a critical examination of each of the following concepts: the libido theory, the Oedipus complex, narcissism, feminine psychology, the death instinct, emphasis upon childhood, transference, culture and neuroses, the ego and the id, anxiety, the super-ego, neurotic guilt feelings, masochism, and psychoanalytic therapy. In the case of each of these topics the Freudian observations and theories are presented objectively and as adequately as space would seem to permit; that which is considered correct is pointed out; that which is considered unfounded and erroneous is criticized; and finally alternate interpretations are offered. These alternate interpretations are in terms of Dr. Horney's own theory of personality as stated in *The neurotic personality of our time*. Very briefly that theory is as follows: "The relevant factor in the genesis of neuroses is then neither the Oedipus complex nor any kind of infantile pleasure strivings but all those adverse influences which make a child feel helpless and defenseless and which make him conceive the world as potentially menacing. Because of his dread of potential dangers the child must develop certain 'neurotic trends' permitting him to cope with the world with some measure of safety" (p. 9).

As already stated, Dr. Horney set herself an extremely ambitious task. Each of the sixteen chapters could well have been a volume. Hence, it is in order to say a word about the author's qualifications for her task. Undoubtedly she is thoroughly conversant with psychoanalytic literature. Concerning another aspect of her qualifications she says: "In all sincerity I may say that I regard myself qualified to make the criticisms contained in this book, because I consistently applied Freud's theories for a period of over fifteen years." There is much more evidence of an acquaintance with the literature of psychology other than psychoanalysis than was apparent in her former book, but there is yet in *New ways in psychoanalysis* a regrettable detachment from the great stream of psychological thought reaching back to Plato. This detachment is one of the very few earmarks of the cultish aspect of psychoanalysis to be found in Dr. Horney's work.

Dr. Horney is freer from what might be termed a Freudian "complex" than any author I have read. She has the almost unique distinction of writing objectively about Freud and his doctrines. The book offers one slight and unimportant exception to this objectivity which is, nevertheless, interesting. In the discussion of "feminine psychology" occasionally a bit of personal feeling seems to enter. The theory of penis-envy as a central explanation of the feminine character seems to place a very nearly unbearable burden upon the author's objectivity.

The essential contention of the book is that beyond doubt S. Freud is a genius of the first order as manifested by his numerous original observations concerning human nature, but that many of his theories—resting as they do upon 19th century philosophical premises—are untenable. Dr. Horney offers

other (and in her judgment more tenable) theories or interpretations of these facts. In doing so she manifests courage, objectivity, industry, intelligence, open-mindedness, together with no little common sense, all of which one cannot refrain from admiring. Yet one completes the book with the feeling that in the main one set of personal opinions have been substituted for another set, namely, Dr. Horney's for Dr. Freud's. Herein lies the weakness and the strength of the book. It is at bottom a stimulating personal testimony concerning the valuable and the less valuable, the sound and the erroneous in psychoanalytic theory, made by a sincere, experienced, informed and apparently unbiased student. As such it is an extremely valuable document, perhaps more for the ground that it breaks than for the fruit that it produces. Should the book be so fortunate as to stimulate other students, the clinical and the more academic, to undertake a similar task, a more balanced and definitive evaluation of psychoanalytical theories would be greatly hastened.

E. V. PULLIAS.

George Pepperdine College, Los Angeles, California.

PSYCHOLOGY FOR BUSINESS AND INDUSTRY. By Herbert Moore. New York: McGraw-Hill, 1939. Pp. xi+527.

In his *Psychology for business and industry* Professor Moore has excellently synthesized the current "aspects of personnel problems in business and industry to which psychology has given, or can give, a contribution." After opening with an introductory chapter on the general service of psychology to business and industry, the author very logically discusses the successive steps which are encountered from the entrance of a prospective worker into the personnel office to the final sales and advertising decisions of the president of the firm.

The chapter headings demonstrate this orderly arrangement of the material for easy assimilation by college students, to wit: Getting a Job; Importance of Proper Selection; Hiring the Worker; Testing the Applicant; Introducing and Constructing Tests; Job Analysis and Job Specification; Promoting the Employee; Training the Worker; Motivating the Worker; Employee Accidents; Employee Fatigue; The Problem Employee; Psychological Problems in Advertising; and Psychological Problems in Selling.

These chapter headings alone form a very enticing "show case" and indicate the author's plan for the orderly presentation of the latest scientific data. The book is well documented with pertinent investigations and experimental evidence. Each chapter concludes with a useful bibliography of required as well as supplementary readings, plus several provocative Topics for Report, which make excellent written problems.

The very division of the book into 15 chapters makes it especially convenient as a classroom text; and the entire volume is obviously pointed toward use as a basic text with classes who have an alert and practical desire for psychology as an applied science instead of a purely academic subject for dilettantes in the liberal arts.

Because Dr. Moore has presented the most significant aspects of personnel and industrial problems and has summarized the researches on these various subjects, his book forms the latest volume in the field, and should, therefore, lend itself admirably to use as a basic college textbook.

GEORGE W. CRANE.

Northwestern University.

MENTAL THERAPY: STUDIES IN FIFTY CASES. By Louis S. London. New York: Covici-Friede, 1939. 2 vols. Pp. 774.

It would be exceedingly unfair to judge this work according to the expectations set up by the blurb on the dust-cover, a blurb particularly unfortunate in a work whose sale "because of its scientific nature, is positively restricted to physicians, lawyers, ministers, educators, and social workers." *Mental Therapy* will not, as is claimed therein, "displace Havelock Ellis' 'Studies in the Psychology of Sex,'" nor is it likely to become "a landmark in the history of psychoanalysis." (At least authors of scientific books should be held directly responsible for blurbs!)

The material is divided into six sections, the first of which is on metapsychology. In this, very sketchy accounts of the evolution of psychotherapeutics and psychoanalysis are presented, and the meaning of the dream, the "psychosexual psychology of the child," and the "psychosexual pathology of the sexual instinct" are discussed. Aside from these 35 pages of discussion there is little more of basic theorizing. The five other sections are concerned almost entirely with the presentation of case material on neuroses, the paraphiliac neuroses (sexual perversions), the border-line neuroses, the schizophrenic psychoses and the cyclic psychoses.

It was difficult for the reviewer to overcome an impression gained early in the reading of these volumes that a more appropriate title for them would have been: "Jottings from a Notebook on Fifty Cases Treated by Psychotherapeutic Means." The actual title leads one to expect a great deal which is not delivered. The author, a man of apparently extensive experience in the analysis of various types of aberrant states, is probably an excellent therapist. Relatively raw data are not, however, the material for a book, no matter how "fascinating" they may be. The theoretical and historical discussion, as already indicated, is sketchy and the case material presented in an unorganized and repetitious fashion. The author's own theories are not developed in sufficient detail to make an adequate evaluation possible. Essentially a Freudian standpoint is taken.

A most unsatisfactory aspect of the presentation is the fact that information as to the details of the analyses is very meager, thus aiding the reader little towards a real understanding of the therapeutic process. We have gone far beyond the *propaganda* stage in psychopathology, the period when it was necessary to accumulate case records to convince even psychologists that certain types of aberrations exist. What interests us now is their significance for normal psychology, their meaning and how they are to be treated. The author discusses the latter two aspects, but generally in a very incomplete fashion. It is especially disappointing that he has not given more along these lines in the case of the psychoses, for it is with respect to this group of disorders that the psychoanalytic literature is most inadequate and detailed accounts of treatment rare.

The form of presentation also leaves much to be desired. The style is in certain respects too popular and in addition not smooth. Paragraphs frequently consist of single sentences, emphasizing the note-like nature of the work. The citations are few and incompletely given; typographical errors occur not infrequently, and terms are not always used precisely.

It is very difficult to recommend these volumes for much more than an inadequate source book of histories of various kinds of patients on whom a type of Freudian therapy has been carried out. What the reader will generally

get will be some notion of the behavior and symptoms of the patient, dream material with partial interpretations, some interpretation of behavior, and a short statement of outcome. A glossary and index are provided.

DAVID SHAKOW.

Worcester State Hospital.

OUT OF THE RUNNING. By Gertrude Hoopes, with a foreword by Edgar Doll. Springfield, Illinois: Charles C. Thomas, 1939. Pp. 158.

Research students interested in work with crippled children have often wished for a comprehensive description of the mental processes of those who are most seriously handicapped. In *Out of the running* we have material of this type. The author, a case of athetosis from inter-craneal injury, has described in detail some of the experiences of her early youth and of her adolescence. Like all reminiscences, they have the emotional color of adulthood rather than the attitudes and emotional sets of the age which the author is describing. These reminiscences are contained in a chapter on childhood and one on youth. There is a very fine discussion of the religious life of the author from adolescence through maturity. The last few chapters are given over to a discussion of hospitals and institutions for the handicapped. It is refreshing to have a discussion of these institutions from the point of view of the person to whose welfare they contribute.

The book suffers from the fact that the author appears to have received very little psychological training of the technical sort. Had her introspection been undertaken after thorough psychological training, it would have contributed far more to our knowledge of the mental processes of the handicapped. Though reminiscences and self-study of the non-technical type are undoubtedly of interest, they cannot be said to make a contribution of great scientific value.

ADA HART ARLITT.

University of Cincinnati.

A SYMPOSIUM ON THE RELATION BETWEEN THE INDIVIDUAL AND THE GROUP. *The American Journal of Sociology*, 1939, **44**, No. 6. Chicago: University of Chicago Press.

The American Journal of Sociology has devoted an entire issue to the problem of the relation between the individual and the group. It presents twelve contributions from eminent scholars in four fields—four in sociology, five in psychology, two in psychiatry, and one in cultural anthropology.

The contributors, named in the order in which their articles appear, are Florian Znaniecki, Maurice Halbwachs,[1] R. S. Woodworth, William E. Blatz, John E. Anderson, J. F. Brown, Kurt Lewin, Floyd H. Allport, Thomas M. French, Harry Stack Sullivan, Bronislaw Malinowski, and Louis Wirth. The titles of the papers may be omitted without appreciable loss. Some of them are inconveniently long, and all of them are less informative as labels than are the authors' names. Wirth's paper is a discussion of the other papers in this issue. By condensing the different contributions and ranging them side by side, by throwing into relief their essential and distinctive features, and above all by evaluating them from a single point of view, he achieves a measure of retrospective unity for the enterprise as a whole.

[1] The article by Halbwachs has been translated from the French by Professor John H. Mueller of Indiana University.

The immediate impression created by the symposium is akin to that expressed in Bertrand Russell's generalization concerning studies of animal learning. "One may say broadly," Russell asserts, "that all animals that have been carefully observed have behaved so as to confirm the philosophy in which the observer believed before his observations began."[2] Apparently groups and individuals behave in a similar manner as regards the convictions of the student of their relation to each other. But this first impression, though it never quite vanishes, is not the one that becomes dominant. The very fact that the same problem reappears persistently, that it is recognizably the same and yet presents different aspects from different points of view, puts the stamp of reality on the problem itself.

For psychologists, at least for those who study human beings, this fact means that they must seriously consider the possibility that the more subtle, elaborate, and complicated processes are inherent in the problems of psychology from the start, and must be reckoned with throughout. Probably the most direct evidence on this point is that presented by Anderson, Blatz, and Lewin; for despite their differences in theory and practice, all three of these men reveal the young of the human species unmistakably responsive to social influences which are far from simple and obvious. Indeed the question of what is simple and what complex—including the general question of what constitutes "fundamental" research—becomes itself less obvious than many once found it. Woodworth's highly original contribution arouses similar reflections. By calling attention to the rôle of the object in team work and in group participation, he not only makes the stimulus-response formula more inclusive and complicated than ever; he also suggests important and pervasive implications for any psychology which is stated primarily in terms of reactions to objects. In Allport's hypothesis concerning conforming behavior, psychologists will find another recognition of the complexities of human reaction, coupled in Allport's treatment with a method of empirical investigation which employs variables congenial to the tough-minded. But perhaps for their most stimulating exercises in intellectual orientation, psychologists should turn to the contributions that come from outside psychology itself, in the more restricted sense of the term.

The purpose of the editors in assembling these papers is stated in the following sentences. "The diversity of viewpoints between and within the different disciplines seems to be responsible in large measure for the present confusion in the psychological and social sciences. In order to clarify the outstanding issues the editors decided to devote this issue of the *Journal* to a symposium upon 'The Relation between the Individual and the Group.'"[3] The editors do not claim that this clarification has been achieved; at any rate, in their comment on the product they are presenting no mention is made of clarification. Instead they say, "The editors believe that this issue will stimulate a reconsideration of basic points of view which will be significant for further research."[4] The reviewer is in entire agreement with the editors if they mean that the outcome of the symposium is stimulation rather than clarification.

EDNA HEIDBREDER.

Wellesley College.

[2] Russell, B. *An outline of philosophy*. London: Allen and Unwin, 1927, 32.
[3] From "In this Issue," two columns of editorial comment immediately following the Table of Contents.
[4] *Loc. cit.*

NEW BOOKS RECEIVED

BORING, E. G., LANGFELD, H. S., AND WELD, H. P. *Introduction to psychology.* New York: John Wiley & Sons, 1939. Pp. xxii+652. Price $3.00.

FAY, J. W. *American psychology before William James.* New Brunswick, N. J.: Rutgers University Press, 1939. Pp. x+240. Price $2.50.

JOHNSON, M. W. *Verbal influences on children's behavior.* Ann Arbor, Michigan: University of Michigan Press, 1939. Pp. ix+191. Price $1.50.

KAMIAT, A. H. *Social forces in personality stunting.* Cambridge, Mass.: Sci-Art Publishers, 1939. Pp. 256. Price $2.50.

McFARLAND, M. B. *Relationships between young sisters as revealed in their overt responses.* New York: Bureau of Publications, Teachers College, Columbia University, 1938. Pp. xi+230. Price $3.50.

O'HARA, F. J. *Psychology and the nurse.* Philadelphia: W. B. Saunders Co., 1939. Pp. 252. Price $1.75.

ROBERTS, H. *The troubled mind.* New York: E. P. Dutton & Co., 1939. Pp. iv+284. Price $2.00.

TERRY, G. C., AND RENNIE, T. A. C. *Analysis of parergasia.* New York: Nervous and Mental Disease Publishing Co., 1938. Pp. 202. Price $4.00.

WECHSLER, I. S. *A textbook of clinical neurology.* (4th ed. revised.) Philadelphia: W. B. Saunders Co., 1939. Pp. 844. Price $7.00.

Science Library 170 OCT 18 1939

A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

CONTENTS

ARTICLES PAGE

Lengths of Conversations: A Conformity Situation Analyzed by the Telic-Continuum and J-Curve Hypothesis 419
Floyd H. Allport and *Richard S. Solomon*

A Method for Evaluating the Experimental Use of Drugs in Psychopathology 465
Milton M. Parker

Psychological States Resulting from Parathyroid Deficiency 481
Carol Finley Dda

Cognitive Aspects of Deterioration in Mental Patients . . 497
Richard S. Solomon

Recent Applications of The A-S Reaction Study 518
Richard Ruggles and *Gordon W. Allport*

SHORT ARTICLES AND NOTES

Effect of Collective Learning upon the Individual Participants 529
Herbert Gurnee

The Diagnostic Value of the Verbal Summator 533
Mary Adah Trussell

A Note on the Relation between Divergent Interests and Emotional Stability 539
Rudolf Pintner and *George Forlano*

REVIEWS 542

BOOKS RECEIVED 552

Volume 34 OCTOBER, 1939 Number 4

ORIGINAL ARTICLES and editorial communications should be sent to:

Dr. Gordon W. Allport, Emerson Hall
Cambridge, Mass.

BOOKS FOR REVIEW should be sent to:

Dr. Edna Heidbreder, Pendleton Hall
Wellesley College, Wellesley, Mass.

FIFTY REPRINTS are supplied gratis to contributors of articles nine pages or more in length. For shorter contributions and for book reviews, five copies of the Journal are supplied gratis.

THE JOURNAL is issued quarterly in January, April, July and October. The yearly volume averages 560 pages. It begins with the January issue and is concluded with the October issue. The subscription per year is $5.00, foreign $5.25; single numbers $1.50.

SUBSCRIPTIONS, orders, and other business communications should be addressed to:

The Journal of Abnormal and Social Psychology
The Ohio State University
Columbus, Ohio